This Page Is Inserted by IFW Operations and is not a part of the Official Record

BEST AVAILABLE IMAGES

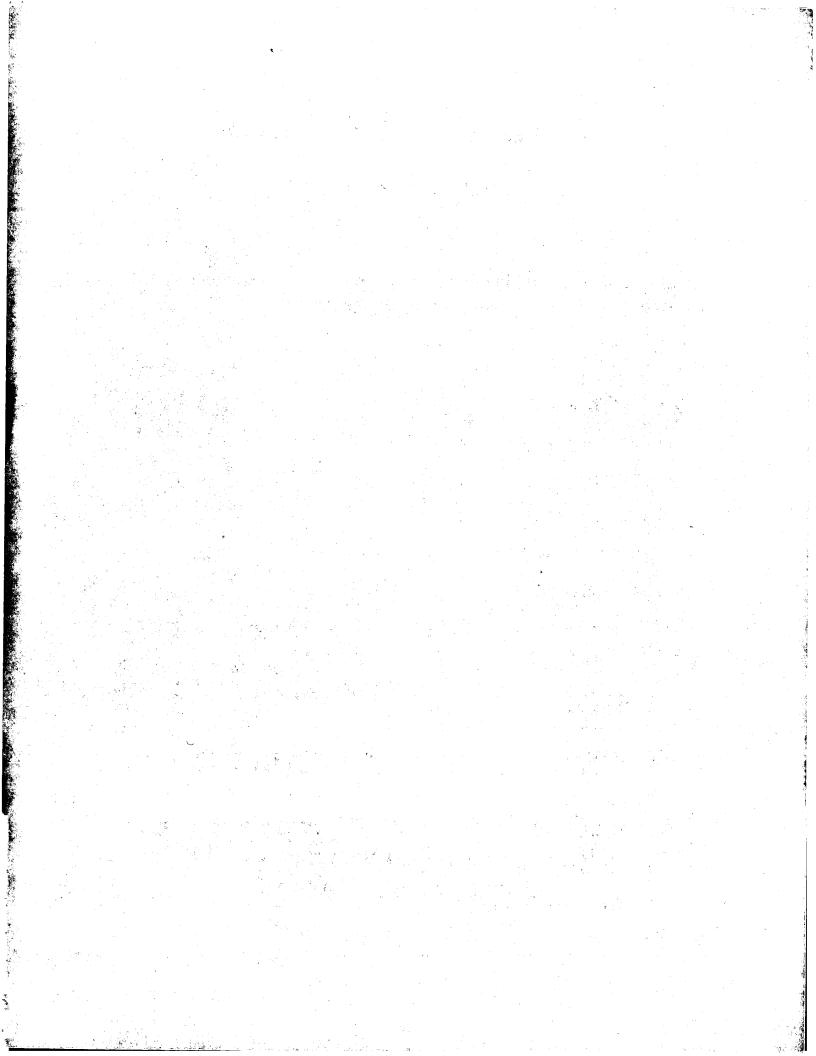
Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents will not correct images, please do not report the images to the Image Problem Mailbox.



PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION International Bureau





INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶: C12N 15/29, C07K 14/415, C12N 5/04, 5/14, A01H 5/00, C12N 15/10, 15/82

(11) International Publication Number:

WO 99/19492

(43) International Publication Date:

22 April 1999 (22.04.99)

(21) International Application Number:

PCT/EP98/06977

(22) International Filing Date:

9 October 1998 (09.10.98)

(30) Priority Data:

PO 9745

10 October 1997 (10.10.97)

ΑU

(71) Applicant (for all designated States except US):
RHONE-POULENC AGRO [FR/FR]; 14/20, rue Pierre
Baizet, F-69009 Lyon (FR).

(72) Inventors; and

- (75) Inventors/Applicants (for US only): DOUTRIAUX, Marie-Pascale [FR/FR]; 64, route de Villebon, F-91160 Saulx les Chartreux (FR). BETZNER, Andreas, Stefan [AU/AU]; 40-Dallachy Place, Page, ACT 2614 (AU). FREYSSINET, Georges [FR/FR]; 21, rue de Nervieux, F-69450 Saint Cyr au Mont d'Or (FR). PEREZ, Pascal [FR/FR]; 17, chemin de la Pradelle, Varennes, F-63450 Chanonat (FR).
- (74) Agent: GENIN, Patrick; Rhône-Poulenc Agro, DPI, 14/20, rue Pierre Baizet, F-69009 Lyon (FR).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published

Without international search report and to be republished upon receipt of that report.

(54) Title: METHODS FOR OBTAINING PLANT VARIETIES

(57) Abstract

An isolated and purified DNA molecule comprising a polynucleotide sequence encoding a polypeptide functionally involved in the DNA mismatch repair system of a plant.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

Ilbania	ES FI FR GA GB GB GH GN GR HU IE IL	Spain Finland France Gabon United Kingdom Georgia Ghana Guinea Greece Hungary Ireland Israel	LS LT LU LV MC MD MG MK	Lesotho Lithuania Luxembourg Latvia Monaco Republic of Moldova Madagascar The former Yugoslav Republic of Macedonia Mali Mongolia	SI SK SN SZ TD TG TJ TM TR TR	Slovenia Slovakia Senegal Swaziland Chad Togo Tajikistan Turkmenistan Turkey Trinidad and Tobago
austria australia azerbaijan kosnia and Herzegovina karbados kelgium kurkina Faso kulgaria kenin	FR GA GB GE GH GN GR HU IE IL	Prance Gabon United Kingdom Georgia Ghana Guinea Greece Hungary Ireland	LU LV MC MD MG MK ML MN	Luxembourg Latvia Monaco Republic of Moldova Madagascar The former Yugoslav Republic of Macedonia Mali	SN SZ TD TG TJ TM TR	Senegal Swaziland Chad Togo Tajikistan Turkmenistan Turkey Trinidad and Tobago
ustralia zerbaijan tosnia and Herzegovina arbados telgium turkina Faso tulgaria tenin	GA GB GB GH GN GR HU IE IL	Gabon United Kingdom Georgia Ghana Guinea Greece Hungary Ireland	LV MC MD MG MK ML MN	Latvia Monaco Republic of Moldova Madagasicar The former Yugoslav Republic of Macedonia Mali	SZ TD TG TJ TM TR TT	Swaziland Chad Togo Tajikistan Turkmenistan Turkey Trinidad and Tobago
zerbaijan Sosnia and Herzegovina Jarbados Jelgium Sulgaria Jenin	GB GE GH GN GR HU IE IL	United Kingdom Georgia Ghana Guinea Greece Hungary Ireland	MC MD MG MK ML MN	Monaco Republic of Moldova Madagascar The former Yugoslav Republic of Macedonia Mali	TD TG TJ TM TR TT	Chad Togo Tajikistan Turkmenistan Turkey Trinidad and Tobago
cosnia and Herzegovina larbados leigium kurkina Faso bulgaria lenin	GB GH GN GR HU IE IL	Georgia Ghana Guinea Gueece Hungary Ireland	MD MG MK ML MN	Republic of Moldova Madagascar The former Yugoslav Republic of Macedonia Mali	TG TJ TM TR TT	Togo Tajikistan Turkmenistan Turkey Trinidad and Tobago
sarbados Jelgium Surkina Faso Jelgaria Jenin Jeazil	GH GN GR HU IE IL	Ghana Guinea Greece Hungary Ireland	MG MK ML MN	Madagascar The former Yugoslav Republic of Macedonia Mali	TJ TM TR TT	Tajikistan Turkmenistan Turkey Trinidad and Tobago
elgium turkina Faso tulgaria tenin trazil	GN GR HU IE IL	Guinea Greece Hungary Ireland	MK ML MN	The former Yugoslav Republic of Macedonia Mali	TM TR TT	Turkmenistan Turkey Trinidad and Tobago
Burkina Faso Bulgaria Benin Brazil	GR HU IE IL	Greece Hungary Ireland	ML MN	Republic of Macedonia Mali	TR TT	Turkey Trinidad and Tobago
Bulgaria Jenin Brazil	HU IE IL	Hungary Ireland	MN	Mali	TT	Trinidad and Tobago
lenin Brazil	IE IL	Ireland	MN	*******		
Frazil	IL			Mongolia	UA	
		Israel				Ukraine
lelarus	10		MR	Mauritania	UG	Uganda
	12	Iceland	MW	Malawi	US	United States of America
Canada	Tl	Italy	MX	Mexico	UZ	Uzbekistan
Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
witzerland	KG	Kyrgyzstan	NO	Norway	zw	Zimbabwe
Côte d'Ivoire	KP	Democratic People's	NZ	New Zealand		
Cameroon		•	PL	Poland		
China	., KR		PT	Portugal		
Cuba		Kazakstan	RO	Romania		
	ıc	Saint Lucia	RU	Russian Federation		
iermany	,	Liechtenstein	SD	Sudan		
Denmark			SR			
istonia						
h h z	te d'Ivoire mercon ina ba ech Republic rmany nmark	te d'Ivoire KP meroon ina KR ba KZ ech Republic LC rmany Li nmark LK	te d'Ivoire KP Democratic People's Republic of Korea meroon KR Republic of Korea ba KZ Kazakstan ech Republic , LC Saint Lucia Liechtenstein mark LK Sri Lanka	te d'Ivoire KP Democratic People's NZ meroon Republic of Korea PL ina KR Republic of Korea PT ba KZ Kazakstan RO ech Republic LC Saint Lucia RU rmany LI Liechtenstein SD nmark LK Sri Lanka SE	te d'Ivoire KP Democratic People's NZ New Zealand meroon Republic of Korea PL Poland ina KR Republic of Korea PT Portugal ba KZ Kazakstan RO Romania ech Republic LC Saint Lucia RU Russian Federation rmany LI Liechtenstein SD Sudan nmark LK Sri Lanka SE Sweden	te d'Ivoire KP Democratic People's NZ New Zealand meroon Republic of Korea PL Poland ma KR Republic of Korea PT Portugal ba KZ Kazakstan RO Romania ech Republic LC Saint Lucia RU Russian Federation rmany LI Liechtenstein SD Sudan mmark LK Sri Lanka SB Sweden

WO 99/19492

10

Methods for Obtaining Plant Varieties

TECHNICAL FIELD

The present invention relates to nucleotide sequences which encode polypeptides involved in the DNA mismatch repair systems of plants, and to the polypeptides encoded by those nucleotide sequences. The invention also relates to nucleotide sequences and polypeptide sequences for use in altering the DNA mismatch repair system in plants. The invention also relates to a process for altering the DNA mismatch repair system of a plant cell, to a process for increasing genetic variations in plants and to processes for obtaining plants having a desired characteristic.

BACKGROUND OF THE INVENTION

Plant breeding essentially relies on and makes use of genetic variation which occurs naturally within and between members of a family, a genus, a species or a subspecies. Another source of genetic variation is the introduction of genes from other organisms which may or may not be related to the host plant.

Allelic loci or non-allelic genes which constitute or contribute to desired quantitative (e.g. growth performance, yield, etc.) or qualitative (e.g. deposition, content and composition of seed storage products; pathogen resistance genes; etc.) traits that are absent, incomplete or inefficient in a species or subspecies of interest are typically introduced by the plant breeder from other species or subspecies, or *de novo*. This introduction is often done by crossing, provided that the species to be crossed are sexually compatible. Other means of introducing genomes, individual chromosomes or genes into plant cells or plants are well known in the art. They include cell fusion, chemically aided transfection (Schocher et al., 1986, Biotechnology 4: 1093) and ballistic (McCabe et al., 1988, Biotechnology 6: 923), microinjection (Neuhaus et al., 1987, TAG 75: 30), electroporation of protoplasts (Chupeau et al., 1989, Biotechnology 7: 53) or microbial transformation methods such as Agrobacterium mediated transformation (Horsch et al., 1985, Science 227: 1229; Hiei et al., 1996, Biotechnology 14: 745).

1.

...

However, when a foreign genome, chromosome or gene is introduced into a plant, it will often segregate in subsequent generations from the genome of the recipient plant or plant cell during mitotic and meiotic cell divisions and, in consequence, become lost from the host plant or plant cell into which it had been introduced. Occasionally, however, the introduced genome, chromosome or gene physically combines entirely or in part with the genome, chromosome or gene of the host plant or plant cell in a process which is called recombination.

Recombination involves the exchange of covalent linkages between DNA molecules in regions of identical or similar sequence. It is referred to here as homologous recombination if donor and recipient DNA are identical or nearly identical (at least 99%)

base sequence identity), and as homeologous recombination if donor and recipient DNA are not identical but are similar (less than 99% base sequence identity).

The ability of two genomes, chromosomes or genes to recombine is known to depend largely on the evolutionary relation between them and thus on the degree of sequence similarity between the two DNA molecules. Whereas homologous recombination is frequently observed during mitosis and meiosis, homeologous recombination is rarely or never seen.

From a breeder's perspective, the limits within which homologous recombination occurs, therefore, define a genetic barrier between species, varieties or lines, in contrast to homeologous recombination which can break this barrier. Homeologous recombination is thus of great importance for plant breeding. Accordingly there is a need for a process for enhancing the frequency of homeologous recombination in plants. In particular, there is a need for a process of increasing homeologous recombination to significantly shorten the length of breeding programs by reducing the number of crosses required to obtain an otherwise rare recombination event.

At least in Escherichia coli, homologous and homeologous recombination are known to share a common pathway that requires among others the proteins RecA, RecB, RecC. RecD and makes use of the SOS induced RuvA and RuvB, respectively. It has been suggested that mating induced recombination follows the Double-Strand Break Repair 20 model (Szostak et al., 1983, Cell 33, 25-35), which is widely used to describe genetic recombination in eukaryotes. Following the alignment of homologous or homeologous DNA double helices the RecA protein mediates an exchange of a single DNA strand from the donor helix to the aligned recipient DNA helix. The incoming strand screens the recipient helix for sequence complementarity, seeking to form a heteroduplex by hydrogen 25 bonding the complementary strand. The displaced homologous or homeologous strand of the recipient helix is guided into the donor helix where it base pairs with its counterpart strand to form a second heteroduplex. The resulting branch point then migrates along the aligned chromosomes thereby elongating and thus stabilising the initial heteroduplexes. Single stranded gaps (if present) are closed by DNA synthesis. The strand cross overs 30 (Holliday junction) are eventually resolved enzymatically to yield the recombination products.

Although in wild type E. coli homologous and homeologous recombination are thus mechanistically similar if not identical, homologous recombination in conjugational crosses E. coli x E. coli occurs five orders of magnitude more frequently than homeologous recombination in conjugational crosses E. coli x S. typhimurium (Matic et al. 1995; Cell 80, 507-515). The imbalance in favour of homologous recombination was shown to be caused largely by the bacterial MisMatch Repair (MMR) system since its

inactivation increased the frequency of homeologous recombination in *E. coli* up to 1000 fold (Rayssiguier et al. 1989, Nature 342, 396-401).

In E. coli, the MMR system (reviewed by Modrich 1991, Annual Rev Genetics 25, 229-253) is composed of only three proteins known as MutS, MutL and MutH. MutS recognizes and binds to base pair mismatches. MutL then forms a stable complex with mismatch bound MutS. This protein complex now activates the MutH intrinsic single stranded endonuclease which nicks the strand containing the misplaced base and thereby prepares the template for DNA repair enzymes.

During recombination, MMR components inhibit homeologous recombination. In vitro experiments demonstrated that MutS in complex with MutL binds to mismatches at the recombination branch point and physically blocks RecA mediated strand exchange and heteroduplex formation (Worth et al., 1994; PNAS 91, 3238-3241). Interestingly, the SOS dependent RuvAB mediated branch migration is insensitive to MutS/MutL, explaining the observed slight increase in SOS dependent homeologous recombination.

15 Homeologous mating even induces the SOS response, thereby taking advantage of RuvAB induction (Matic et al. 1995, Cell 80, 507-515).

The MMR system thus appears to be a genetic guardian over genome stability in E. coli. In this role it essentially determines the extent to which genetic isolation, that is, speciation, occurs. The diminished sensitivity of the SOS system to MMR, however, allows (within limits) for rapid genomic changes at times of stress, providing the means for fast adaptation to altered environmental conditions and thus contributing to intraspecies genetic variation and species evolution.

1

14... 1.4

The important role of MMR in preserving genomic integrity has been established also in certain eukaryotes. In its efficiency, the human MMR, for example, may even counteract potential gene therapy tools such as triple-helix forming oligonucleotides including RNA-DNA hybrid molecules (Havre et al., 1993, J. Virology 67: 7234-7331; Wang et al., 1995, Mol. Cell. Biol. 15: 1759-1768; Kotani et al., 1996, Mol. Gen. Genetics 250: 626-634; Cole-Strauss et al., 1996, Science 273: 1387-1389). Such oligonucleotides are designed to introduce single base changes into selected DNA target sequences in order to inactivate for example cancer genes or to restore their normal function. The resulting base mismatches however are recognised by the mismatch repair system which then directs removal of the mismatched base, thereby reducing the efficiency of oligonucleotide induced site-specific mutagenesis.

To date, two families of related genes, homologous to the bacterial MutS and MutL genes have been identified or isolated in yeast and mammals (recent reviews by Arnheim and Shibata, 1997, Curr. Opinion Genet. Dev. 7, 364-370; Modrich and Lahue, 1996, Annual Rev. Biochem. 65, 101-133; Umar and Kunkel, 1996, Eur. J. Biochem. 238, 297-307). Biochemical and genetic analysis indicated that eukaryotic MutS homologs (MSH)

WO 99/19492 4 PCT/EP98/06977

and MutL homologs (MLH, PMS), respectively, fulfil similar protein functions as their bacterial counterparts. Their relative abundance, however, could reflect different mismatch specificity and/or specialisation for different tissues or organelles or developmental processes such as mitotic versus meiotic recombination.

To date, six different genes homologous to *MutS* have been isolated in yeast (yMSH), and their homologs have been found in mouse (mMSH) and human (hMSH), respectively. Encoded proteins yMSH2, yMSH3 and yMSH6 appear to be the main MutS homologs involved in MMR during mitosis and meiosis in yeast, where the complementary proteins MSH3 and MSH6 alternatively associate with MSH2 to recognise different mismatch substrates (Masischky et al., 1996, Genes Dev. 10, 407-420). Similar protein interactions have been demonstrated for the human homologs hMSH2, hMSH3 and hMSH6 (Acharya et al., 1996, PNAS 93, 13629-13634).

MutL homologs (MLH and PMS), recently reviewed by Modrich and Lahue (1996, Annual Rev. Biochem. 65. 101-133) have so far been found in yeast (yMLH1 and yPMS1), mouse (mPMS2) and human (hMLH1, hPMS1 and hPMS2). The hPMS2 is a member of a family of at least 7 genes (Horii et al., 1994, Biochem. Biophys. Res. Commun. 204, 1257-1264) and its gene product is most closely related to yPMS1. Prolla et al. (1994, Science 265, 1091-1093) presented evidence for yPMS1 and yMLH1 to physically associate with each other and, together, to interact with the MutS homolog yMSH2 to form a ternary complex involved in mismatch substrate binding.

However, while medical interest in mismatch repair has prompted extensive research on MMR in bacteria, yeast and mammals, MMR genes have not been isolated from higher plants prior to the present invention and no attempts to adjust the plant MMR to plant breeding needs have been reported.

SUMMARY OF THE INVENTION

According to a first embodiment of the invention, there is provided an isolated and purified DNA molecule comprising a polynucleotide sequence encoding a polypeptide functionally involved in the DNA mismatch repair system of a plant. In one form of this embodiment, the invention provides an isolated and purified DNA molecule comprising a polynucleotide sequence encoding a polypeptide which is homologous to a mismatch repair polypeptide of a yeast or of a human. More particularly, the invention provides polynucleotide sequences encoding polypeptides which are homologous to the mismatch repair polypeptides MSH3 and MSH6 of Saccharomyces cerevisiae. Still more particularly, the invention provides the coding sequences of the genes AtMSH3 and AtMSH6 of Arabidopsis thaliana. as defined hereinbelow, and polynucleotide sequences encoding polypeptides which are homologous to polypeptides encoded by AtMSH3 and AtMSH6.

ä

蓝

£.....

4 . . .

According to a second embodiment of the invention, there is provided an isolated and purified polypeptide functionally involved in the DNA mismatch repair system of a plant, for example a polypeptide which is homologous to a mismatch repair polypeptide of a yeast or of a human such as a polypeptide encoded by the genes AtMSH3 or AtMSH6 of 5 Arabidopsis thaliana, as defined hereinbelow.

According to a third embodiment of the invention, there is provided an isolated and purified DNA molecule comprising a polynucleotide sequence selected from the group consisting of (i) a sequence encoding a polynucleotide which is capable of interfering with the expression of a plant polynucleotide sequence encoding a polypeptide which is homologous to a mismatch repair polypeptide of a yeast or of a human and thereby disabling said plant polynucleotide sequence; and (ii) a sequence encoding a polypeptide capable of disrupting the DNA mismatch repair system of a plant.

According to a fourth embodiment of the invention there is provided a chimeric gene comprising a DNA sequence selected from the group consisting of (i) a sequence encoding a polynucleotide which is capable of interfering with the expression of a plant polynucleotide sequence encoding a polypeptide which is homologous to a mismatch repair polypeptide of a yeast or of a human and thereby disabling said plant polynucleotide sequence, and (ii) a sequence encoding a polypeptide capable of disrupting the DNA mismatch repair system of a plant: together with at least one regulation element capable of functioning in a plant cell. Examples of such regulation elements include constitutive, inducible, tissue type specific and cell type specific promoters such as 35S. NOS, PR1a. AoPR1 and DMC1. Typically, a chimeric gene of the fourth embodiment will also include at least one terminator sequence, more typically exactly one terminator sequence.

In the third and fourth embodiments, said interference, by said polynucleotide sequence, with the expression of a plant polynucleotide sequence encoding a polypeptide which is homologous to a mismatch repair peptide of a yeast or a human typically occurs by hybridisation or by co-suppression.

According to a fifth embodiment of the invention there is provided a plasmid or vector comprising a chimeric gene of the fourth embodiment. A vector of the fifth 30 embodiment may be, for example, a viral vector or a bacterial vector.

According to a sixth embodiment of the invention, there is provided a plant cell stably transformed, transfected or electroporated with a plasmid or vector of the fifth embodiment.

According to seventh embodiment of the invention, there is provided a plant as comprising a cell of the sixth embodiment.

According to an eighth embodiment of the invention, there is provided a process for at least partially inactivating a DNA mismatch repair system of a plant cell, comprising

WO 99/19492 PCT/EP98/06977

transforming or transfecting said plant cell with a DNA sequence of the third embodiment or a chimeric gene of the fourth embodiment or a plasmid or vector of the fifth embodiment, and causing said DNA sequence to express said polynucleotide or said polypeptide.

According to a ninth embodiment of the invention, there is provided a process for increasing genetic variation in a plant comprising obtaining a hybrid plant from a first plant and a second plant, or cells thereof, said first and second plants being genetically different; altering the mismatch repair system in said hybrid plant; permitting said hybrid plant to self-fertilise and produce offspring plants; and screening said offspring plants for plants in which homeologous recombination has occurred. For example, homeologous recombination may be evidenced by new genetic linkage of a desired characteristic trait or of a gene which contributes to a desired characteristic trait.

According to a tenth embodiment of the invention there is provided a process for obtaining a plant having a desired characteristic, comprising altering the mismatch repair system in a plant, cell or plurality of cells of a plant which does not have said desired characteristic, permitting mutations to persist in said cells to produce mutated plant cells, deriving plants from said mutated plant cells, and screening said plants for a plant having said desired characteristic.

In a preferred form of the ninth and tenth embodiments of the invention, the step of altering the mismatch repair system comprises introducing into said hybrid plant, plant, cell or cells a chimeric gene of the fourth embodiment and permitting the chimeric gene to express a polynucleotide which is capable of interfering with the expression of a plant polynucleotide sequence in a mismatch repair gene of the hybrid plant, plant, cell or cells, or a polypeptide capable of disrupting the DNA mismatch repair system of the hybrid plant or cells.

In other embodiments, the invention provides (a) an oligonucleotide capable of hybridising at 45°C under standard PCR conditions to a DNA molecule of the first embodiment; (b) an oligonucleotide capable of hybridising at 45°C under standard PCR conditions to the DNA of SEQ ID NO: 18 and (c) an oligonucleotide capable of hybridising at 45°C under standard PCR conditions to the DNA of SEQ ID NO:30; with the proviso that the oligonucleotide of (a), (b) and (c) is other than SEQ ID NO:1 or SEQ ID NO:2.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 provides a diagrammatic representation of the primer sequences used to 35 isolate AtMSH3.

Figure 2 is a plasmid map of clone 52, showing restriction enzyme cleavage sites in the 5' half of the full-length cDNA for AtMSH3.

ate.

- 45 ·

11

1

Figure 3 is a plasmid map of clone 13, showing restriction enzyme cleavage sites in the 3' half of the full-length cDNA for AtMSH3.

Figure 4 is a sequence listing of the coding sequence of AtMSH3, together with a deduced sequence of the encoded polypeptide.

Figure 5 is a protein alignment of yeast (Saccharomyces cerevisiae) and Arabidopsis thaliana MSH3 protein.

Figure 6 provides a diagrammatic representation of the primer sequences used to isolate AtMSH6.

Figure 7 is a plasmid map of clone 43, showing restriction enzyme cleavage sites in the 5' half of the full-length cDNA for AtMSH6.

Figure 8 is a plasmid map of clone 62, showing restriction enzyme cleavage sites in the 3' half of the full-length cDNA for AtMSH6.

Figure 9 is a sequence listing of the coding sequence of AtMSH6, together with a deduced sequence of the encoded polypeptide.

Figure 10 is a protein alignment of yeast (Saccharomyces cerevisiae) and Arabidopsis thaliana MSH6 protein.

Figure 11 is a genomic sequence listing of AtMSH6.

Figure 12 is a plasmid map of plasmid pPF13.

Figure 13 is a plasmid map of plasmid pPF14.

Figure 14 is a plasmid map of plasmid pCW186.

Figure 15 is a plasmid map of plasmid pCW187.

Figure 16 is a plasmid map of plasmid pPF66.

Figure 17 is a plasmid map of plasmid pPF57.

Figure 18 is a diagrammatic representation of an antisense gene construction for use in homeologous meiotic recombination.

Figure 19 is a plasmid map of plasmid p3243.

DETAILED DESCRIPTION OF THE INVENTION

The present invention is based on the inventors' discovery that there exist in higher plants genes which are homologous to MMR genes in E. coli, and to MMR genes in yeasts and humans.

Thus, the inventors have identified genes, herein designated AtMSH3 and AtMSH6, of the plant Arabidopsis thaliana which encode the proteins AtMSH3 and AtMSH6. These plant proteins are homologous to yMSH3 and yMSH6, respectively. The present inventors have isolated cDNAs encoding the proteins AtMSH3 and AtMSH6 and have isolated the complete gene encoding AtMSH6. Given the teaching herein, other genes (for example AtMSH2, and genes of other plants) may be obtained which are involved in DNA mismatch repair in plants, including other genes which encode polypeptides homologous to MMR proteins of yeasts or humans, such as genes which encode

WO 99/19492 PCT/EP98/06977

polypeptides homologous to yeast MSH2. MLH1 or PMS2, or to human MLH1. PMS1 or PMS2. For example, given the teaching herein, genes of members of the *Brassicaceae* tamily or of other unrelated tamilies, for example the *Poaceae*, the *Solanaceae*, the *Asteraceae*, the *Malvaceae*, the *Fabaceae*, the *Linaceae*, the *Canabinaceae*, the *Dauaceae* and the *Cucurbitaceae* family, and which encode polypeptides homologous to MMR proteins of yeasts or humans may be obtained.

Examples of plants whose genes encoding polypeptides homologous to MMR proteins of yeasts or humans may be obtained given the teaching herein include maize, wheat, oats, barley, rice, tomato, potato, tobacco, capsicum, sunflower, lettuce, artichoke, safflower, cotton, okra, beans of many kinds including soybean, peas, melon, squash, cucumber, oilseed rape, broccoli, cauliflower, cabbage, flax, hemp, hops and carrot.

Within the meaning of the present invention, a first polypeptide is defined as homologous to a second polypeptide if the amino acid sequence of the first polypeptide exhibits a similarity of at least 50% on the polypeptide level to the amino acid sequence of the second polypeptide.

A procedure which may be followed to obtain genes AtMSH3 and AtMSH6 is described in Example 1. Essentially the same technique may be applied to obtain other mismatch repair genes of Arabidopsis thaliana, and essentially the same technique as 20 exemplified herein may be applied to cDNA obtained by reverse transcription of RNA from other plants. Alternatively, given the sequence information disclosed herein, other degenerate oligonucleotide primers, especially oligonucleotides of the invention which are capable of hybridising at 45°C under standard PCR conditions (such as the conditions described in Example 1 using primers UPMU and DOMU) to AtMSH3 and/or AtMSH6 25 may be designed and obtained for use in isolating sequences of plant mismatch repair genes which are homologous to AtMSH3 or AtMSH6, from other plants. oligonucleotides of the invention which are capable of hybridising at 45°C under standard PCR conditions to plant mismatch repair genes of plants other than Arabidopsis thaliana also fall within the scope of the present invention and may be utilised to obtain mismatch Typically, such oligonucleotides are capable of 30 repair genes of still other plants. hybridising at 45°C under standard PCR conditions to a DNA molecule which encodes a polypeptide which is homologous to a mismatch repair polypeptide of a yeast or a human. The temperature at which oligonucleotides of the invention hybridise to AtMSH3 and/or AtMSH6, or to plant mismatch repair genes of plants other than Arabidopsis thaliana, or 35 to DNA molecules which encode polypeptides which are homologous to a mismatch repair polypeptide of a yeast or a human may be higher than 45°C, for example at least 50°C, or G at least 55°C, or at least 60°C or as high as 65°C.

The successful gene isolation disclosed herein demonstrates for the first time the existence of MMR in higher plants and indicates the presence of other plant MMR genes. For example, genes encoding the plant homologs of MSH1, MSH2, MSH4, MSH5, PMS1, PMS2 and MLH1 may be identified given the teaching herein. Such genes, as well as those specifically described herein, separately or in combination, are useful in manipulating the plant MMR for plant breeding purposes. Thus, for example, the plant MMR may be altered by including in a plant cell a polynucleotide sequence as defined herein above with reference to the third embodiment of the invention, and causing the polynucleotide sequence to express either a polynucleotide which disables a plant MMR gene, or a polypeptide which disrupts the plant's MMR system.

The DNA molecule of the third embodiment of the invention includes a polynucleotide sequence (herein referred to as a MMR altering gene) which may for example encode sense, antisense or ribozyme molecules characterised by sufficient base sequence similarity or complementarity to the gene to be altered to permit the antisense or 15 ribozyme molecule to hybridise with the plant MMR gene in vivo or to permit the sense molecule to participate in co-suppression. Alternatively, the MMR altering gene may encode a protein or proteins which interfere with the activity of a plant MMR protein and thus disrupt the plant's MMR system. For example, such encoded proteins may be antibodies or other proteins capable of interfering with MMR protein function, such as by 20 complexing with a protein functionally involved in plant MMR thereby disrupting the MMR of the plant. An example of such a protein is the MSH3 protein of Arabidopsis thaliana described herein or a protein of another plant which is homologous to the MSH3 protein of A. thaliana. For instance, overexpression of MSH3 in a plant cell causes MSH2 present in the cell to be substantially completely complexed, disrupting the 25 mismatch repair mechanism or mechanisms in the cell which are functionally dependent on the presence of a complex of MSH2 with MSH6. Similarly, mismatch repair mechanisms which depend on the presence of a complex of MSH2 and MSH3 may be disrupted by the overexpression of MSH6.

A chimeric gene of the fourth embodiment, incorporating a MMR altering gene, may be prepared by methods which are known in the art. Similarly, the MMR altering gene may be introduced into a plant cell, regenerating tissue or whole plant by techniques known in the art as being suitable for plant transformation, or by crossing. Known transformation techniques include Agrobacterium tumefaciens or A. rhizogenes mediated gene transfer, ballistic and chemical methods, and electroporation of protoplasts.

The MMR altering gene or genes are typically expressed from suitable promoters. Suitable promoters may direct constitutive expression, such as the 35S or the NOS promoter. Usually, however, the promoter will direct either inducible or tissue specific (e.g. callus; embryonic (issue; etc.), cell type specific (e.g. protoplasts; meiocytes; etc.) or developmental (e.g. embryo) expression of the altering gene or genes. in order for the

MMR system to function in tissue types or cell types, or at developmental stages of the plant, in which it is not desirable for the MMR system to be altered. Using such promoters, therefore, the activity of a MMR altering gene may be limited to a specific stage during plant development or it may be altered by controlling conditions external to the plant, and the deleterious effects of a permanently disabled or altered DNA mismatch repair system in a plant may be avoided. Examples of suitable promoters which are not constitutive are known in the art and include inducible promoters such as PR1a (reviewed by Gatz, 1997, Annual Rev. Plant Phys. Plant Mol. Biol. 48: 89), tissue specific promoters such as AoPR1 (Sabahattin et al., 1993, Biotechnology 11: 218), and cell-type specific promoters such as DMC1.

A chimeric gene in accordance with the invention may further be physically linked to one or more selection markers such as genes which confer phenotypic traits such as herbicide resistance, antibiotic resistance or disease resistance, or which confer some other recognisable trait such as male sterility, male fertility, grain size, colour, growth rate, flowering time, ripening time, etc.

The process of the tenth embodiment of the invention provides, for example, a process for generating intraspecies genetic variation by altering the mismatch repair system in a plant cell, in regenerating plant tissue or in a whole plant. The plant cell, regenerating tissue or whole plant includes and expresses one or more MMR altering 20 genes which are capable of altering mismatch repair in the plant cell, regenerating tissue or whole plant. Alteration of MMR may be achieved, for example, by inactivating the genes encoding plant MSH3 and/or plant MSH6. It is preferred to inactivate the plant MSH3 and MSH6 encoding genes at the same time and in the same plant cell, regenerating tissue or whole plant. Typically in this preferred form of the invention 25 inactivation of either plant MSH3 or MSH6 alone is insufficient to substantially alter the plant's mismatch repair system and only when both MSH3 and MSH6 are inactivated simultaneously is the plant's mismatch repair system sufficiently altered to prevent the MMR system from recognising base pair mismatches, base insertions or deletions as a result of DNA replication errors. DNA damage, or oligonucleotide induced site-specific 30 mutagenesis. However, in some applications of the invention, inactivation of only one gene may also be used to cause genomic instability or increase the efficiency of sitespecific mutagenesis.

If desired, the MMR altering gene or genes may later be rendered non-functional or ineffective, or may be removed from the genome of the plant cell, regenerating tissue or whole plant in order to restore mismatch repair in the plant cell, regenerating tissue or whole plant. The MMR altering gene or genes may be inactivated by means of known gene inactivation tools, such as ribozymes, or may be removed from the genome using gene elimination systems known in the art, such as CRE/LOX. It is preferred to render two genes, whose gene products combine to incapacitate MMR, ineffective by separating

the altering genes through segregation. Therefore, in a preferred embodiment of the invention a first plant cell or plant is generated in which only plant MSH3 is incapacitated, and a second plant cell or plant is generated in which only plant MSH6 is incapacitated. The combination of both genomes, for example by crossing, then produces significant MMR deficiency in those cells or plants which have inherited both altering genes. If the altering genes are expressed from unlinked loci, gene segregation restores MMR activity in the progeny of the cells or plants.

In a process of the ninth embodiment of this invention, homeologous recombination is enhanced between different genomes, chromosomes or genes in plant cells or plants by altering MMR in said plant cells or plants. Such genomes, chromosomes or genes are characterised in that they originate from different plant families, genera, species, subspecies, plant varieties or lines. Hybrid plant cells or hybrid plants may be produced by crossing, by cell fusion or by other techniques known in the art. These plant cells or plants are further characterised by expressing one or more genes that are capable of altering mismatch repair in the plant cell or plants.

In the process of the ninth embodiment, the homeologous recombination is typically for the purpose of introducing a desired characteristic in the hybrid plant. In this typical application of the process of the ninth embodiment, and in the process of the tenth embodiment the desired characteristic may be any characteristic which is of value to the plant breeder. Examples of such characteristics are well known in the art and include altered composition or quality of leaf or seed derived storage products (e.g. oil, starch, protein), altered composition or quality of cell walls (e.g. decrease in lignin content), altered growth rate, prolonged flowering, increased plant yield or grain yield, altered plant morphology, resistence to pathogens, tolerance to or improved performance under environmental stresses of various kinds, etc.

In a preferred form of the tenth embodiment, an MMR altering gene is cointroduced along with the homeologous genome, chromosome or gene of another plant
cell or plant into an MMR proficient plant cell or MMR proficient plant to produce a
hybrid plant cell or hybrid plant in which homeologous recombination can occur.

Suitably, the MMR proficient plant cell or MMR proficient plant may also include an
MMR altering gene. For example a gene capable of inactivating plant MSH3 may be cointroduced along with the homeologous genome, chromosome or gene of another plant
cell or plant into an MMR proficient plant cell or MMR proficient plant in which MSH6
is inactivated. A resultant hybrid plant in which homeologous recombination occurs will
include both the MSH3 and MSH6 altering genes and its MMR system will therefore be
inactivated.

In this form of the invention, if hybrid plants are to be produced by crossing, the MMR altering gene preferably originates from the male parent, thus ensuring that the

MMR altering gene is always introduced and is not present in the recipient cell. That is, the MMR of the recipient cell, prior to introduction of the MMR altering gene, is typically proficient. Alternatively, if an MMR altering gene is present in a recipient cell it may be ineffective or inefficient on its own, or it may be linked to an inducible or tissue specific or cell type specific promoter which only renders the MMR altering gene active under limited conditions.

Thus, in a preferred form of the process of the ninth embodiment, the MMR system of the hybrid plant is initially unaltered. In this form of the process, the step of altering the mismatch repair system may comprise introducing into the hybrid plant, or cells thereof, a MMR altering gene, such as by Agrobacterium tumefaciens or A. rhizogenes mediated gene transfer, ballistic and chemical methods, and electroporation of protoplasts.

The MMR altering gene or genes are typically expressed from suitable promoters, as described above. Preferably, the promoter is transcriptionally active in mitotically and meiotically active tissue and/or cells to ensure MMR alteration after chromosome pairing at mitosis and meiosis, respectively. The preferred timing for MMR alteration is at meiosis, because recombinant genomes, chromosomes or genes are directly transmitted to the progeny. A suitable meiocyte specific promoter is for example the DMC1 promoter from Arabidopsis thaliana ssp. Ler. (Klimyuk and Jones, 1997, Plant J. 11, 1-14). However, mitotic homeologous recombination is also a desirable outcome as somatic recombination events can be transmitted to offspring due to the totipotency of plant cells and the lack of predetermined germ cells in plants.

If desired, the MMR altering gene or genes may later be rendered non-functional or ineffective, or may be removed from the hybrid plant or hybrid plant cells, in order to restore mismatch repair in the hybrid plant or hybrid plant cells. The MMR altering gene or genes may be inactivated by means of known gene inactivation tools as described herein above.

EXAMPLES

Example 1. Cloning of the AtMSH3 and AtMSH6 coding sequences

Isolation of partial AtMSH3 and AtMSH6 consensus sequences

Degenerate oligonucleotides UPMU (SEQ ID NO:1) and DOMU (SEQ ID NO:2)

UPMU CTGGATCCACIGGICCIAA(C/T)ATG

DOMU CTGGATCC(A/G)TA(A/G)TGIGTI(A/G)C(A/G)AA

were used to isolate AtMSH3 and AtMSH6 sequences by PCR amplification.

Primers UPMU and DOMU correspond to conserved amino acid sequences of the proteins MutS (E. coli and S. typhimurium), HexA (S. pneumoniae). Repl (mouse) and Ducl (human). The conserved regions to which they are targeted are TGPNM for UPMU (amino acid positions 852-856 for AtMSH6 and 816-820 for AtMSH3) FATHY or FVTHY

· 3

for DOMU (amino acid positions 964-968 for AtMSH6 and 928-932 for AtMSH3, respectively.) These primers have been used to isolate MSH2 and MSH1 from yeast (Reenan and Kolodner, Genetics 132: 963-973 (1992)) and MSH2 from *Xenopus* and mouse (Varlet et al., Nuc. Acids Res. 22:5723-5728 (1994)).

Template single strand cDNA was produced by reverse transcription of 2 µg total RNA from a cell suspension culture of Arabidopsis thaliana ecotype Columbia (Axelos et al. 1989, Mol. Gen. Genetics 219: 106-112). The PCR reaction was performed under the following conditions in a final volume of 100µl: 0.2mM dNTP, 1µM each primer, 1XPCR buffer, 1u Taq DNA polymerase (Appligene) in the presence of template cDNA. PCR 10 parameters were 5 minutes at 94°C, followed by 30 cycles of 40 seconds at 95°C, 90 seconds at 45°C, 1 minute at 72°C. The amplification products were cloned into pGEM-T vector (Promega) and sequenced. Two different clones were isolated, S5 (350bp) was homologous to MSH3, S8 (327bp) was homologous to MSH6. Complete cDNA sequences were then isolated according to the Marathon cDNA amplification kit procedure (Clontech). 15 In summary, this procedure involves producing double stranded cDNA by reverse transcription of 2µg polyA+ RNA from the cell suspension culture of Arabidopsis, Adaptors are ligated on each side of the cDNA. The ligated cDNA is used as a template for 5° and 3° RACE PCR reactions in the presence of primers that are specific for the adaptor on one side (AP1 and AP2), and specific for the targeted gene on the other side. A 5% and a 3° 20 fragment that overlap are thus produced for each gene. The complete gene coding sequence can be reconstituted taking advantage of a unique restriction site, if available, in the overlapping region. Specific details of this procedure as it was used to isolate AtMSH3 and .41MSH6 coding regions, are as follows.

Isolation of AtMSH3 complete coding sequence

From the sequence of clone S5, primer 636 (SEQ ID NO:3) was designed:

TGCTAGTGCCTCTTGCAAGCTCAT.

Primer AP1 (SEQ ID NO:4) is complementary to a portion of an adaptor sequence which had been ligated to the 5' and 3' ends of Arabidopsis cDNA:

API CCATCCTAATACGACTCACTATAGGGC.

PCR performed on the ligated cDNA with primers 636 and AP1 for the 5' RACE PCR was followed by a second round of amplification with the nested primers AP2 (SEQ ID NO:5) and S525 (SEQ ID NO:6)

AP2 ACTCACTATAGGGCTCGAGCGGC

S525 AGGTTCTGATTATGTGTGACGCTTTACTTA

35 (the latter was also designed to correspond to a part of the sequence of clone S5) and produced a 2720bp DNA fragment. Figure 1 provides a diagrammatic representation of the primer sequences used to isolate AtMSH3. Another primer (S51, SEQ ID NO:7)

۲,

S51 GGATCGGGTACTGGGTTTTGAGTGTGAGG

was designed closer to the 5' border and permitted the determination of 99bp upstream to the ATG initiation codon. For the 3' RACE PCR, a first PCR reaction was performed with primers AP1 and 635 (SEQ ID NO:8).

635 GCACGTGCTTGATGGTGTTTTCAC

- 5 followed by a second round of amplification, using the nested primers AP2 and S523 (SEQ ID NO:9)
- S523 TCAGACAGTATCCAGCATGGCAGAAGTA
 which produced a DNA fragment of 890bp. Both DNA fragments were subcloned into
 pGEM-T and sequenced. Since PCR amplification using the Expand Long Template PCR
 system (Boehringer-Mannheim) produced errors in the sequence, new oligonucleotides
 were designed to isolate those sequences again by PCR, but with the high fidelity DNA
 polymerase Pfu. PCR with primers 1S5 (SEQ ID NO:10) and S53 (SEQ ID NO:11)
 - 1S5 ATCCCGGGATGGGCAAGCAAAAGCAGCAGACGA
 - S53 GACAAAGAGCGAAATGAGGCCCCTTGG
- amplified the 1244bp fragment clone 52 (SEQ ID NO:12, cloned into pUC18/Sma1). PCR with primers S52 (SEQ ID NO:13) and 2S5 (SEQ ID NO:14)
 - 2S5 ATCCCGGGTCAAAATGAACAAGTTGGTTTTAGTC
 - S52 GCCACATCTGACTGTTCAAGCCCTCGC
- amplified the 2104bp clone 13 (SEQ ID NO:15, cloned into pUC18/Sma1). The complete coding sequence of the AtMSH3 gene was reconstructed in pUC18 by ligating the 5' half of AtMSH3 (clone 52) to the 3' half of AtMSH3 (clone 13) after digesting with BamH1 which has a unique cleavage site in the overlapping region of both clones. This manipulation yielded plasmid pPF26. The Smal fragment from pPF26 contains the complete AtMSH3 coding sequence. The remaining primers referred to in Figure 1 are as follows:
 - S51 GGATCGGGTACTGGGTTTTGAGTGTGAGG (SEQ ID NO:16)
 - S525 AGGTTCTGATTATGTGTGACGCTTTACTTA (SEQ ID NO:17)

Figures 2 and 3 provide plasmid maps of clones 52 and 13 respectively, showing restriction enzyme cleavage sites. The complete AtMSH3 coding sequence (SEQ ID NO:18) 30 is 3246bp long and is shown in Figure 4 together with the deduced sequence (SEQ ID NO:19) of the encoded polypeptide. AtMSH3 is clearly homologous to the yeast and mouse MSH3 genes. A sequence alignment of polypeptides encoded by AtMSH3 and that encoded by Saccharomyces cerevisiae MSH3 is set out in Figure 5.

Isolation of the AtMSH6 complete coding sequence and genomic sequences

The same procedure allowed isolation of the AtMSH6 cDNA. Figure 6 provides a diagrammatic representation of the primer sequences used to isolate AtMSH6. For the 5' RACE PCR, primers 638 (SEQ ID NO:20) and AP1 (SEQ ID NO:4)

638 TCTCTACCAGGTGACGAAAAACCG
allowed the amplification of a 2889 DNA fragment. Primer S81 (SEQ ID NO:21)

- 14

4,000

S81 CGTCGCCTTTAGCATCCCCTTCCTTCAC

helped define the 142bp upstream to the ATG initiation codon. On the 3' side, RACE PCR was initially performed with primers \$823 (SEQ ID NO:22) and API (SEQ ID NO:4).

S823 GCTTGGCGCATCTAATAGAATCATGACAGG

5 and then with the nested primers 637 (SEQ ID NO:23) and AP2 (SEQ ID NO:5).

637 GACAGCGTCAGTTCTTCAGAATGC

to produce a 774bp DNA fragment. As for AtMSH3, those fragments were cloned and sequenced. Re-isolation of the DNA sequence using the high fidelity Pfu polymerase and newly designed primers 1S8 (SEQ ID NO:24) and S83 (SEQ ID NO:25) (for the 5' side) led to a 2182 bp DNA fragment identified as clone 43 (SEQ ID NO:26, cloned in pUC18/Sma1), and a 1379bp clone identified as clone 62 (SEQ ID NO:27, also cloned in pUC18/Sma1).

158	ATCCCGGGATGCAGCGCCAGAGATCGATTTTGT	
2S8	ATCCCGGGTTATTTGGGAACACAGTAAGAGGATT (SEQ	ID
NO:28)		
S82	GCGTTCGATCATCAGCCTCTGTGTTGC (SEQ ID NO:29)	•
S83	CGCTATCTATGGCTGCTTCGAATTGAG	

Figures 7 and 8 provide plasmid maps of clones 43 and 62 respectively, showing restriction enzyme cleavage sites. Clones 43 and 62 were digested by the Xmn1 restriction enzyme for which a unique site is present in their overlapping region and then ligated. The complete AtMSH6 coding sequence (SEQ ID NO:30) is 3330bp long and is shown in Figure 9 together with the deduced sequence (SEQ ID NO:31) of the encoded polypeptide. AtMSH6 is clearly homologous to the yeast and mouse MSH6genes. A sequence alignment of polypeptides encoded by AtMSH6 and that encoded by Saccharomyces cerevisiae MSH6 is set out in Figure 10.

An AtMSH6 genomic sequence was also isolated from a genomic DNA library constituted after partial Sau3AI digestion of DNA from the Arabidopsis cell suspension. 8062bp were sequenced that covered the AtMSH6 gene and show colinearity with the cDNA. 16 introns are found scattered along the gene. The complete genomic sequence 30 (SEQ ID NO:98) is shown in Figure 11.

Example 2. A measure of somatic variation in MMR deficient plants

Constructs

15

Constructs with antisense AtMSH3 or antisense AtMSH6 or both AtMSH3/AtMSH6 under the control of a single 35S promoter have been inserted into the binary vector pPZP121 (Hajdukiewicz et al., Plant Mol. Biol. 23, 793-799) between the right and left borders of the T-DNA. The pPZP121 plasmid confers chloramphenicol resistance to Escherichia coli or Agrobacterium tumefaciens bacteria. The aacC1 gene is carried by the T-DNA and allows selection of transformed plant cells on gentamycin (Hajdukiewicz et al., Plant Mol. Biol. 25, 989-994). For the purpose of expressing antisense constructs, a 35S

promoter/terminator cassette with a polylinker was introduced into pPZP121. The 3' ends of the considered genes have been chosen since this region seems more efficient for antisense inhibition. For AtMSH3 this corresponds to clone 13 (2104bp), for AtMSH6 this corresponds to clone 62 (1379bp). Clone 13 comprises 2104bp of the 3' region that were cut off the pUC18 vector by Sal1/Sst1 restriction, blunted with T4 DNA polymerase and ligated into the T4 DNA polymerase blunted BamHI site of pPZP121/35S, creating clone pPF13. The same procedure was followed for the 3' region of AtMSH6 clone 62 (1379bp) thus creating plasmid pPF14. For the double constructs, the 3' region (from clone 62) of AtMSH6 was introduced ahead of the AtMSH3 region into pPF13 creating pCW186 and reciprocally, the 3' region of AtMSH3 (from clone 13) was introduced ahead of AtMSH6 into pPF14, creating pCW187.

These constructs were introduced into the Arabidopsis cells (as described below) of wildtype Columbia and of the Columbia tester line.

An alternative strategy to antisense inhibition of AtMSH6 comes from experiments of Marra et al. (1998, Proc. Natl. Acad. Sci USA 95, 8568-8573) who show that overexpression of functional MSH3 results in depletion of MSH6 protein in human cells. This depletion may generate a mismatch repair mutant phenotype.

For the purpose of overexpressing functional AtMSH3 protein in plant cells, the complete MSH3 coding region was excised from pPF26 (example 1) by digestion with 20 Smal, and was inserted into the Smal site of pCW164. The resulting construct was named pPF66. It contains a complete AtMSH3 gene under the control of the 35S promoter inside the left (LB) and right (RB) border of the T-DNA. This T-DNA also contains the hpt2 gene for gentamycin selection. Plasmid pPF66 was introduced into Arabidopsis cells as described below. One cell clone was selected which clearly overexpressed the AtMSH3 gene as shown by Northern analysis. Figures 12-16 provide plasmid maps of plasmids pPF13, pPF14, pCW186, pCW187 and pPF66, respectively.

Construction of tester construct

For the purpose of Forward Mutagenesis Assays, a tester construct was built containing the coding regions for nptII, codA, uidA. All three genes are driven by the 35S promoter and are terminated by the 35S terminator. This construct was obtained by introducing an EcoR1 fragment encoding the codA cassette (2.5kb) and a HindIII fragment encoding the uidA (GUS) cassette (2.4kb) into the pPZP111 vector (Hajdukiewicz et al.,1994, Plant Mol Biol 23: 793-799) which already contained the nptII expression cassette. This new plasmid was named pPF57. NptII is used to select for transformed plant cells.

35 GUS is used to analyse the degree of gene silencing in the construct (i.e. to identify cell lines in which the transgenes are expressed), and codA is used as a marker for forward mutagenesis (described below).

٤.

The plasmid map of pPF57 is provided in Figure 17.

Plant cell transformation

then transformed by co-cultivation. A suspension culture of Arabidopsis thaliana cells that 5 has been established by Axelos et al. (1992, Plant Physiol. Biochem. 30, 1-6) may be used. One day old freshly subcultured cells are diluted five times into AT medium (Gamborg B5 medium. 30g/l sucrose, 200µg/l NAA). 10µl of saturated Agrobacterium containing the transforming T-DNA constructs are added to 10ml diluted cells in a 100ml erlenmeyer. The co-cultivation is agitated slowly (80rpm) for 2 days. The cells are then washed 3 times into AT medium and finally resuspended in the same initial volume (10ml). The culture is agitated for 3 days to allow expression before plating on selection plates (AT/BactoAgar 8g/l+gentamycin 50µg/ml). Transformed individual calli are isolated 3 weeks later.

Tester Strain

The tester construct on plasmid pPF57 was introduced into Arabidopsis cells of wildtype Columbia using the transformation protocol described above. Among 10 candidate transformants, one cell clone was shown (by Southern analysis) to have a unique T2DNA insertion. All three genes were shown to be functional in this cell line as indicated by resistance to kanamycin, blue staining in the presence of X-Glu (GUS), and sensitivity to 5-fluoro-cytosine (codA).

MMR altering genes (described above) were then introduced individually into the tester line and transformed cells are used for analysis of both Microsatellite Instability and Forward Mutagenesis.

Microsatellite analysis

Microsatellites have been described in Arabidopsis (Bell and Ecker, 1994, Genomics 19, 137-144). The present Example is based on a study of instability of microsatellites that are polymorphic amongst different ecotypes. DNA is extracted from the transformed calli. Specific primers have been defined that are used to amplify the microsatellite sequence. One of the two primers is previously P³² labelled by T4 kinase. In case of a polymorphic variation, new PCR products appear that do not follow the expected pattern of migration on a polyacrylamide gel. This is a commonly observed feature for MMR deficient cells in yeast or mammalian cells.

In particular, the present Example describes a study on microsatellites ca72 (CA₁₈), nga172 (GA₂₉), and ATHGENEA(A₃₉), chosen because they belong to the types predominantly affected in human mismatch repair deficient tumors. The size of these microsatellites is not conserved from one *Arabidopsis* ecotype to the other.

Arabidopsis cells which are transformed with an MMR altering gene (above) and control cells not expressing the MMR altering gene are allowed to form calli. DNA is

rapidly extracted from the calli and is analysed for microsatellite instability as described in detail by Bell and Ecker 1994. Genomics 19, 137-144. In summary, the relevant microsatellite is amplified by PCR using P32 labelled primers. The PCR products are separated on a DNA sequencing gel for size determination. Size differences between microsatellites from transformed and control cells not expressing the MMR altering gene in question indicate microsatellite instability as a result of MMR alteration.

The sequences of primers used for PCR amplification of microsatellites ca72 and nga172 are included in Table 1. PCR amplification of microsatellite ATHGENEA made use of a forward primer containing the sequence

ACCATGCATAGCTTAAACTTCTTG (SEQ ID NO:32)

and of a reverse primer containing the sequence

ACATAACCACAAATAGGGGTGC (SEQ ID NO:33).

The amplification for microsatellite ca72 revealed in Columbia control cells (with respect to the MMR altering gene) a 248 bp long PCR fragment instead of the published length of 124 bp. DNA sequencing verified this fragment as a CA₁₈ microsatellite.

Forward mutagenesis assav

10

Tester cells transformed with antisense AtMSH3 or antisense AtMSH6 or both AtMSH3/AtMSH6 are analysed for the stability of the codA gene. The functional codA gene confers to sensitivity to 5-fluoro-cytosine (5FC), whereas a gene inactivating mutation in 20 codA will confer resistance to 5FC. The frequency of resistant cells is therefore a good indicator of somatic variation as a direct result of MMR alteration. Variants resistant to 5FC are first analysed for GUS activity. If GUS is inactive, 5FC resistance is assumed to be due to gene silencing (all three genes are under the 35S promoter). If GUS is active, 5FC resistance is assumed to be due to forward mutations that have inactivated codA. PCR is then performed on the putative codA mutant genes which is then sequenced to confirm the presence of forward mutations in codA.

Besides codA, other marker genes may also be used for the Forward Mutagenesis Assay such as the ALS gene (conferring sensitivity to valine or to sulfonylurea; Hervieu and Vaucheret, 1996, Mol. Gen. Genet. 251 220-224; Mazur et al. 1987, Plant Physiol. 85 1110-30 1117).

Example 3. Homeologous meiotic recombination in Arabidopsis thaliana

A. Construction of a meiocyte specific gene expression cassette comprising the DMC1 promoter and the NOS terminator

(i) The DMC1 promoter may be used as published by Klimyuk and Jones, 1997, 35 Plant J. 11.1-14). To obtain a more convenient alternative for gene cloning, a 3.3 Kb

ij

10

20

long subfragment of the DMC1 promoter was obtained by PCR from genomic DNA of Arabidopsis thaliana (ssp. Landsberg erecta "Ler").

The PCR was done in three rounds:

Round One: A 3.7 Kb long product was obtained using the forward primer 5 DMCIN-A comprising the sequence

GAAGCGATATTGTTCGTG (SEQ ID NO:34)

and the reverse primer DMCIN-B comprising the sequence

AGATTGCGAGAACATTCC (SEQ ID NO:35).

The weak amplification product was then used as template for round two and three.

Round Two: A 3.1 Kb long product comprising the promoter and the 5' untranslated leader was obtained using forward primer DMCIN-1, which contained the sequence

acgcgtcgacTCAGCTATGAGATTACTCGTG (SEQ ID NO:36)

and introduced a Sall cloning site at the 5' end of the promoter fragment, and reverse primer DMCIN-2 which contained the sequence

gctctagaTTTCTCGCTCTAAGACTCTCT (SEQ ID NO:37)

and introduced a XbaI site at the 3° end of the PCR fragment.

Round Three: A 0.2 Kb long product comprising the first exon/intron of the *DMC*1 promoter was obtained using forward primer DMCIN-3, which contained the sequence

gctctagaGCTTCTCTTAAGTAAGTGATTGAT (SEQ ID NO:38)

and introduced a XbaI site at the 5' end of the PCR fragment, and reverse primer DMCIN-4, containing the sequence

teccegggetegagagatetecatggTTTCTTCAGCTCTATGAATCC (SEQ ID NO:39) and introduced at the 3' end of the PCR product restriction sites for Ncol. Bg/II. Xhol and 25 Smal.

The products obtained in round Two and Three were digested with XbaI and subsequently ligated to reconstitute a 3.3 Kb long DMC1 promoter from which the first two in-frame ATG start codons were replaced with a unique restriction site for XbaI. This promoter can be cloned between the restriction sites for SaII and SmaI of p3264, which contains the SacI-EcoRI NOS terminator in pBIN19, to yield the entire expression cassette in pBIN19. This cassette contains the following cloning sites: NcoI, BgIII, XhoI. SmaI and (already present on p3264) KpnI and SacI.

(ii) Another strategy yielded the following convenient *DMC*1 promoter. A 1.8 kb DNA fragment comprising the 3' terminal part of the meiocyte specific *DMC*1 promoter was isolated by PCR from purified genomic DNA of *Arabidopsis thaliana* (ssp. Landsberg erecta "Ler"). The forward PCR primer (DMC1a) contained the sequence

acgcgtcgacGAATTCGCAAGTGGGG (SEQ ID NO:40)

and introduced a Sast cloning site at the 5' end of the promoter fragment. The reverse PCR primer (DMC1b) contained the sequence

introducing a polylinker region at the 3' end of the promoter fragment. The PCR reaction was carried out with VENT DNA Polymerase (NEB) over 25 cycles using the following cycling protocol: 1 minute at 94°C, 1 minute at 54°C, 2 minutes at 72°C.

The PCR reaction yielded a blunt ended DNA fragment which was digested with restriction enzyme Sall and was cloned into the cleavage sites of restriction enzymes Sall and Smal in plasmid p2030, a pUC118 derivative containing the Sacl-EcoRl NOS terminator fragment of pBIN121. The cloning yielded plasmid p2031, containing the DMC1 promoter-polylinker-NOS terminator expression cassette depicted in Figure 18.

10 B. Construction of an MSH3 antisense gene under the control of the DMC1 promoter

A 2.1 kb DNA fragment encoding the carboxyterminal part of AtMSH3 was removed from the polylinker of clone 13 described in Example 1 after (i) digestion with KpnI. (ii) blunting of the DNA ends generated by KpnI and (iii) digestion with BamHI. The isolated fragment was then cloned in antisense orientation downstream of the DMC1 promoter in plasmid p2031, which had been digested with restriction enzymes SmaI and BglII. This cloning yielded plasmid p2033 (Figure 18).

After digestion of p2033 with EcoRI, a 4.1 kb DNA fragment was recovered comprising the DMC1 promoter, the partial MSH3 cDNA in antisense orientation with respect to the promoter and the NOS terminator. This fragment was cloned into the EcoRI restriction site of plant transformation vector pNOS-Hyg-SCV to yield plasmid p3242 (Figure 18).

C. Construction of a combined MSH6/MSH3 antisense gene under the control of a single DMC1 promoter

A 3.1 kb fragment, encoding in antisense orientation the partial AtMSH6 and AtMSH3 sequences and the 35S terminator, was isolated from pCW186 by digestion with KpnI. The fragment was treated with Klenow enzyme to blunt both ends. It was then cloned into the SmaI site of plasmid p3243 (a pNOS-Hyg-SCV derivative, illustrated in Figure 19), which had been opened to delete the region between the SmaI sites. Clones containing the fragment in the antisense orientation with respect to the DMC1 promoter (described in A(ii) above) were identified by diagnostic digestion with BamHI. The selected construct was named p3261.

Another practical way of cloning the double antisense gene is as follows. A 1 kb DNA fragment encoding the carboxyterminal part of AtMSH6 is isolated from clone 62 described in Example 1 after digestion of clone 62 plasmid DNA with BamHI, which cleaves in the 5' polylinker region flanking the partial cDNA, and with EcoRI, which cleaves within the cDNA. The isolated fragment is treated with Klenow enzyme to blunt both its ends and is cloned into the recipient plasmid p2033 or p3242. For the purpose of

cloning, the recipient plasmid may be cleaved with either AvaI or NcoI and can be blunted with Klenow enzyme to produce blunt acceptor ends for fragment cloning. This cloning yields two opposite orientations of cloned fragment DNA with respect to the DMC1 promoter. These can be identified by diagnostic digestion with restriction enzymes ScaI or XmnI in conjunction with SacI. The selected construct contains the DMC1 promoter, the combined partial cDNAs for AtMSH3 and AtMSH6 (both cloned in antisense orientation with respect to the DMC1 promoter) and the NOS terminator. If the recipient plasmid is p2033, the combined antisense gene under control the single DMC1 promoter is recovered from the vector after EcoRI digestion and is cloned into the EcoRI restriction site of pNOS-Hyg-SCV.

D. Construction of a full-length MSH3 sense gene under control of the DMC1 promoter for overexpression of functional MSH3 protein

Overexpression of MSH3 protein was shown in human cells (Marra et al., 1998, Proc. Natl. Acad. Sci. USA 95, 8568-8573) to complex all available MSH2 protein. This leaves MSH6 protein without its partner, leading to the degradation of MSH6 protein and eventually to a mismatch repair phenotype.

This phenomenon is exploited to increase homeologous meiotic recombination in Arabidopsis as an alternative to antisense inhibition of AtMSH6. For this purpose the full-length cDNA encoding AtMSH3 is isolated from plasmid pPF66 by digestion with Smal and is cloned into the Smal site of the DMC1 expression cassettes described in A(i).

ž.

...

E. Selection of Recombination markers on homeologous chromosomes of Arabidopsis thaliana subspecies Landsberg erecta (Ler), Columbia (Col) and C24, respectively

E(i). Visual recombination markers in Arabidopsis th. subspecies C24:

Agrobacterium mediated transformation with a T-DNA containing a 35S-GUS gene, inactivated by insertion of a 35S-Ac transposable element (Finnegan et al., 1993, Plant Mol. Biol. 22, 625-633), had yielded a C24 line in which the T-DNA construct was integrated into chromosome 2. Genetic and molecular analysis of this line shows that the Ac transposon had excised from its T-DNA locus thereby restoring GUS activity, but had re-inserted into the chromosome at a distance of 16.4 cM, where it stayed fixed (due to disablement of Ac) within the chlorina gene. Insertional inactivation of the chlorina gene caused a bleached phenotype in those plants that were homozygous for this mutation. Because of the two linked phenotypic markers, chlorina3A:Ac and GUS, this C24 line was used in crosses to wildtype Ler for analysis of meiotic homeologous recombination on chromosome 2 in conjunction with molecular recombination markers.

35 E(ii). Visual recombination markers in Arabidopsis th. Ler:

The Ler line NW1 (obtained from NASC, Nottingham, UK) contains one recessive visual marker per chromosome. i.e. an-1 on Chr.1, py-1 on Chr.2, gll-1 on Chr.3, cer2-1

on Chr.4. and ms1-1 on Chr.5. This line is used in crosses to wildtype C24 which expresses an MMR altering gene for analysis of meiotic homeologous recombination on chromosomes 1-5 in conjunction with molecular recombination markers listed in Table 1.

Other Ler lines from NASC have several visual markers in close proximity to each other on the same chromosome. When these lines are used for hybrid production, analysis of homeologous meiotic recombination may make use entirely of visual recombination markers. Particularly suitable for crossing to C24 wildtype that is expressing a MMR altering gene are the following Ler lines:

NW22: relative markers are dis1 - (4 cM) - ga4 - (11 cM) - th1 on chromosome 1.

NW10: relevant markers are tz-201 - (5 cM) - cer3 on chromosome 5.

NW134, relevant markers are ttg - (4 cM) - ga3 on chromosome 5.

NW24 (abi3-1) and NW64 (gl1-1). When present in the same plant on chromosome 3. abi3-1 and gl1-1 are 13 cM apart. Since this marker combination is not available from NASC, we have combined these markers by crossing line NW24 to line NW64. The F1 offspring were allowed to self-fertilise and to produce F2 seeds. F2 Plants which carry both markers as homozygous traits on the same chromosome can be identified firstly, by germinating F2 seeds on germination medium containing selective concentrations of abscisic acid, and subsequently, by identifying among the abscisic acid resistant plants those individuals which show the glabra phenotype.

20 E(iii) Molecular recombination markers in Col, Ler and C24:

The genome of Arabidopsis thaliana is interspersed with unique base sequences arranged as simple tandem repeats. Allelic repeats can vary in length between different Arabidopsis subspecies and when amplified by PCR yield diagnostic DNA products of different length named Simple Sequence Length Polymorphisms (SSLPs). Many SSLPs have been genetically mapped and have been assigned to unique chromosome locations on the recombinant inbred map (Bell and Ecker, 1994, Genomics 19, 137-144; Lister and Deans lines, Weeds World 4i, May 1997).

In Table 1 are listed 28 mapped and established SSLPs between Ler and Col. A number of PCR primer pairs are described herein (SEQ ID NO:42 to SEQ ID NO:97) which also yielded SSLPs between C24 and Ler (19 SSLPs) or between C24 and Col (25 SSLPs), respectively. Polymorphic SSLPs can be used as molecular markers in the analysis of homeologous recombination between genomes from these subspecies.

The PCR reactions (25 µL) were carried out over 35 cycles (15 seconds at 94°C, 30 seconds at 55°C and 30 seconds at 72°C), with 0.25 U Taq DNA polymerase and 0.6 µg genomic DNA in reaction buffer containing 2 mM MgCl₂. PCR products were separated by agarose gel electrophoresis (4% ultra high resolution agarose) and visualised by ethidiumbromide staining. The results from the PCR experiments are summarised in

Table 1, which also shows the sequence of PCR primers, primer annealing temperature (Tm), PCR product length and chromosome location of SSLP (with respect to the RI map of May 1997, Weeds World 4i).

F. Production of hybrid plants

C24 plants heterozygous for chlorina3A:Ac/GUS are crossed as male to emasculated wildtype Ler to produce Ler/C24(chlorina3A, GUS) hybrid seeds.

Due to the heterozygosity of the C24 parent, only 50 % of hybrid plants have inherited the chlorina3A:Ac/GUS locus. The remaining 50% of hybrid plants are wildtype with respect to chlorina3A:Ac/GUS. Since the mutant locus is linked to a kanamycin resistance gene (contained on the same T-DNA as GUS) mutant plants can be pre-selected by germinating hybrid seeds on germination medium containing 50 mg/L kanamycin.

Ler plants homozygous for the five chromosome markers are male sterile (ms1-1) and are crossed without emasculation to wildtype C24 to produce Ler(an-1, py-1, gl1-1, cer2-1, ms1-1)/C24 hybrid seeds.

Other Ler plants, which are male fertile, are crossed after emasculation of the female parent to produce Ler/C24 hybrid seeds.

G. Introduction of MSH3 and MSH6/3 antisense genes into Arabidopsis and analysis of meiotic homeologous recombination

(i) Transformation of hybrid plants and analysis of homeologous meiotic recombination

The plant transformation vectors comprising the antisense genes described in (B) and (C) above are introduced into Agrobacterium tumefaciens strain AGL1 (Lazo et al., 1991, Bio/Technology 9, 963-967) by electroporation. Recombinant Agrobacterium clones are selected on LB medium containing 50 mg/L rifampicin and 100 mg/L carbenicillin. Selected clones are used to infect roots of Arabidopsis hybrid plants (described in (F) above) using the root transformation protocol of Valvekens et al. (1988, PNAS 85, 5536-5540) except that shoot and root inducing media contain hygromycin (10 mg/L) instead of kanamycin.

Plants regenerated from roots of hybrid plants are genetic clones of root donating plants and therefore are again genetic hybrids of two Arabidopsis subspecies described in (F). However, in contrast to the root donating plants, the regenerated hybrid plants also contain the introduced transgene and the co-introduced hygromycin resistance gene with the latter allowing these plants to grow on hygromycin containing culture medium.

Hygromycin resistant plants are then allowed to enter the reproductive phase and to produce gametes by meiotic divisions of microspore and megaspore mothercells. At meiosis, the *DMC*1 promoter is activated and can direct the expression of antisense genes described in (B) and (C) above, leading to decreased *MSH*6 and/or *MSH*3 gene

expression. This in turn depletes the gamete mothercells of MSH6 and/or MSH3 protein, thus causing alteration of MMR during meiotic divisions and increasing the recombination frequency between homeologous chromosomes.

Transgenic plants are then allowed to self-fertilise and to produce seeds. These seeds (F2 seeds with respect to hybrid production), and the plants derived therefrom, carry the homeologous recombination events which can be identified by using the visual and molecular recombination markers described in (E) above.

In case of homeologous recombination between chromosomes of Ler and C24(chlorina3A:Ac, GUS), the analysis concentrates on chromosome 2 by selecting plants showing the visual phenotypic marker chlorina. This marker thus serves as a reference point as it indicates that respective chromosomes 2 originate from C24. Other markers, such as GUS or molecular markers, on chromosome 2 may then be used to identify chromosomal regions which are derived from the Ler chromosome as a result of homeologous recombination. F2 plants of control transformants not expressing the antisense gene(s) can be analysed in parallel and the results can be used for comparison to homeologous recombination results obtained in antisense plants.

(ii) <u>Transformation of C24 wildtype</u>, hybrid plant production and analysis of homeologous meiotic recombination

Introduction of MMR altering genes into wildtype C24 is done using the root transformation protocol as described in G(i) for transformation of hybrid plants. Transformed plants are selected by resistance to either 10 mg/L hygromycin (in case of transformation with T-DNA's derived from pNOS-Hyg-SCV) or to 50 mg/L kanamycin (in case of transformation with T-DNA's derived from pBIN19).

Transgenic plants are then allowed to self-fertilise and to produce seeds (T1 seeds).

Segregation of the antibiotic resistance gene in the T1 population then indicates the number of transgene loci. Lines with a single transgene locus (indicated by a 3:1 ratio of resistant:sensitive plants) are selected and are bred to homozygosity. This is done by collecting selfed seeds (T2) from T1 plants and subsequent testing of at least four independent T2 seed populations for segregation of the antibiotic resistance gene. T2 populations which do not segregate the antibiotic resistance gene are assumed to be homozygous for both the resistance gene and the linked MMR altering gene.

C24 plants homozygous for the MMR altering gene are then crossed to Ler lines homozygous for recessive visual markers (see E(ii)) to produce C24/Ler hybrid plants as described in (F). These F1 hybrids are then allowed to enter the reproductive phase and to produce gametes by meiotic division of microspore and megaspore mothercells. At meiosis, the DMC 1 promoter is activated and can direct the expression of antisense or sense genes described in (B), (C) and (D) above, leading to decreased MSH6 and/or MSH3 gene expression. This in turn depletes the gamete mothercells of MSH6 and/or MSH3

protein, thus causing alteration of MMR during meiotic divisions and increasing the recombination frequency between the homeologous chromosomes of C24 and Ler. Recombination events are then scored in the F2 generation.

For recombination analysis, the hybrid plants are allowed to self-fertilise and to produce F2 seeds. F2 plants are then preselected for a first visual marker. Since this marker is recessive, its visual presence indicates homozygosity for Ler DNA at the relevant locus. Those F2 plants which show this first visual marker are then analysed for the presence or absence of a second visual marker which in the Ler parent is closely linked to the first marker. Absence of the second visual marker indicates recombination between the relevant C24 and Ler chromosomes between the first and second marker. The frequency of recombination in transgenic hybrids is compared to the recombination frequency in control hybrids not expressing the MMR altering gene.

Examples of recombination analysis are the following.

The Ler line NW22(dis1. ga4. th1) is used for crosses to transformedC24.

F2 plants are preselected first for thiamine requirement (th1) and then are further analysed for re-appearance of wildtype height (loss of ga4) and/or re-appearance of normal trichomes (loss of dis1) as a result of recombination.

The Ler line NW10(tz-201, cer3) is used for crosses to transformedC24.

F2 plants are then preselected first for thiazole requirement (tz) and then are further analysed for re-appearance of normal, i.e. non-shiny stems (loss of cer3) as a result of recombination.

The Ler line NW134 (ttg, ga3) is used for crosses to transformed C24. F2 plants are first preselected for dwarfish appearance (ga3) and are then analysed for re-appearance of trichomes (loss of ttg) as a result of recombination.

Ler plants homozygous for abi3-1 and gl1-1 are used for crosses to transformed C24. F2 plants are first preselected for their ability to germinate in the presence of abscisic acid and are then analysed for re-appearance of trichomes on the leaves (loss of gl1-1) as a result of recombination.

In the case of homeologous recombination between transformed C24 and the Ler line NW1 (an-1, py-1, gll-1, cer2-1, msl-1), recombination analysis is similar the one described above, except that the second marker is not a visual marker but has to be a molecular marker. This is because the Ler parent carries only one visual marker per chromosome.

		TABLE 1: S	ABLE 1: SSLP Markers in Arabidopsis thaliana Subspecies	haliana	Subspecies		
Chromosome	RI Map Position	PCR Primer Pair	Primer Sequence	Tm (°C	length/COL	length/LER	length/C24
_	2.3	ATEATI F	GCCACTGCGTGAATGATATG	57.8	172	162	162
		ATEAT! R	CGAACAGCCAACATTAATTCCC	58.2			
_	9.3	NGA63 F	AACCAAGGCACAGAAGCG	57.3	Ξ	68	120
		NGA63 R	ACCCAAGTGATCGCCACC	59.6			
_	40.1	NGA248 F	TACCGAACCAAAAGG	56.1	143	129	no amplific.
		NGA248 R	TCTGTATCTCGGTGAATTCTCC	58.2			
	81.2	NGA128 F	GGTCTGTTGATGTCGTAAGTCG	60.1	180	061	no amplific.
		NGA128 R	ATCTTGAAACCTTTAGGGAGGG	58.2			
		•					
	81.2	NGA280 F	CTGATCTCACGGACAATAGTGC	1.09	105	85	85
		NGA280 R	GGCTCCATAAAAGTGCACC	87.8			
	111.4	NGAIII F	CTCCAGTTGGAAGCTAAAGGG	09	128	162	170
		NGAIII R	TGTTTTTAGGACAAATGGCG	0/			
=	ca. 7.5	NGA 168 F	CCTICACATCCAAAACCCAC	57.8	213	217	208
		NGA 168 R	GCACATACCCACAACCAGAA	87.8			

1	ca 48	NGA1126L	CGCTACGCTTTTCGGTAAAG	57.8	161	661	961
		NGA1126R	GCACAGTCCAAGTCACAACC	59.9			
=	62.2	NGA361L	AAAGAGATGAGAATTTGGAC	51.7	114	120	114
		NGA361R	ACATATCAATATATTAAAGTAGC	49.5			
	73	NGA168 F	TCGTCTACTGCACTGCCG	59.6	151	135	135
		NGA168 R	GAGGACATGTATAGGAGCCTCG	6.19			
	ca. 77	AihBIO2 L	TGACCTCCTCTTCCATGGAG	59.9	141	209	139
		AthBIO2 R	TTAACAGAAACCCAAAGCTTTC	54.5			
1	ca. 83	AthUBIQUE L	AGGCAAATGTCCATTTCATTG	54.1	146	148	148
		AthUBIQUE R	ACGACATGCCAGATTTCTCC	57.8			
	3.4	NGA172 F	AGCTGCTTCCTTATAGCGTCC	09	162	136	140
		NGA172 R	CATCCGAATGCCATTGTTC	55.4			
	12.8	NGA126 F	GAAAAACGCTACTTTCGTGG	56.1	611	147	no amplific.
		NGA126 R	CAAGAGCAATATCAAGAGCAGC	58.2			
=	17.5	NGA 162 F	CATGCAATTTGCATCTGAGG	55.8	107	89	no amplitic.
		NGA162 R	CICTGTCACTCTTTCCTCTGG	1.09			
				•			

=	8.18	NGA6 F	TGGATTTCTTCCTCTTCAC	56.1	143	123	143
		NGA6 R	ATGGAGAAGCTTACACTGATC	1.95			
١٨	8.61	NGA12 F	AATGTTGTCCTCCCTCCTC	59.9	247	234	220
		NGA12 R	TGATGCTCTCTGAAACAAGAGC	58.2			
	_						
١٨	24.1	NGA8 F	GAGGCAAATCTTTATTTCGG	56.1	154	861	190
		NGA8 R	TGGCTTTCGTTTATAAACATCC	54.5			
	·						
\	102	NGA1107 L	GCGAAAAACAAAAAATCCA	50.2	150	140	140
.4		NGA1107 R	CGACGAATCGACAGAATTAGG	58			
>	8.11	NGA225 F	GAAATCCAAATCCCAGAGGG	58	611	189	611
		NGA225 R	TCTCCCCACTAGTTTTGTGTCC	1.09			
۸	16.7	NGA249 F	TACCGTCAATTTCATCGCC	55.4	125	115	115
		NGA249 R	GGATCCCTAACTGTAAAATCCC	58.2			
>	19.9	CA72 F	AATCCCAGTAACCAAACACACA	56.3	124	110	110
		CA72 R	CCCAGTCTAACCACGACCAC	61.9			
,>	20	NGA151 F	GTTTTGGGAAG1TTTGCTGG	55.8	150	120	130
		NGA151 R	CAGTCTAAAAGCGAGAGTATGATG	58.6			

	,	NCA 10K B	GTTATGGAGITTICTAGGGCACG	1.00	157	123	130
>	55	NG A DA	TGCCCCATTTTGTTCTTCTC	55.8			
		3 00 00					
>	37 R	NGA139 F	AGAGCTACCAGATCCGATGG	59.9	174	132	132
-	2:15	NGA139 R	GGTTTCGTTTCACTATCCAGG	55.8			
>	\$0	NGA76 F	GGAGAAATGTCACTCTCCACC	1.09	231	> 250	300
	3	NGA76 R	AGGCATGGGAGACATTTACG	57.8			
	1 19	ATHSO191 L	ATHSO191 L CTCCACCAATCATG	55.8	148	156	146
•		ATHSO191 R	TGATGTTGATGGAGATGGTCA	53.7			
>	81.7	NGA 129 F	TCAGGAGGAACTAAAGTGAGGG	1.09	177	179	172
		NGA129 R	CACACTGAAGATGGTCTTGAGG	1.09			

CLAIMS

- 1. An isolated and purified DNA molecule comprising a polynucleotide sequence encoding a polypeptide functionally involved in the DNA mismatch repair system of a plant.
- 5 2. A DNA molecule according to claim 1 wherein said polypeptide is homologous to a mismatch repair polypeptide of a yeast or of a human.
 - 3. A DNA molecule according to claim 1 wherein said polypeptide is homologous to AtMSH3 (SEQ ID NO: 19) or to AtMSH6 (SEQ ID NO: 31).
- 4. An isolated and purified polypeptide functionally involved in the DNA no mismatch repair system of a plant.
 - 5. A polypeptide according to claim 4 which is homologous to a mismatch repair polypeptide of a yeast or of a human.
- 6. An isolated and purified polypeptide selected from the group consisting of a polypeptide encoded by the gene AtMSH3 (SEQ ID NO: 18), a polypeptide encoded by the gene AtMSH6 (SEQ ID NO:30), polypeptides homologous to a polypeptide encoded by the gene AtMSH3 (SEQ ID NO: 18) and polypeptides homologous to a polypeptide encoded by the gene AtMSH6 (SEQ ID NO:30).
- 7. An isolated and purified DNA molecule comprising a polynucleotide sequence selected from the group consisting of (i) a sequence encoding a polynucleotide which is capable of interfering with the expression of a plant polynucleotide sequence encoding a polypeptide which is homologous to a mismatch repair polypeptide of a yeast or of a human and thereby disabling said plant polynucleotide sequence; and (ii) a sequence encoding a polypeptide capable of disrupting the DNA mismatch repair system of a plant.
- 8. A DNA molecule according to claim 7 comprising a polynucleotide sequence 25 encoding a polynucleotide capable of interfering with the expression of a plant polynucleotide sequence encoding a polypeptide which is homologous to a mismatch repair polypeptide of a yeast or of a human and thereby disabling said plant polynucleotide sequence.
- 9. A DNA molecule according to claim 8 wherein said polynucleotide is capable 30 of interfering with the expression of a plant polynucleotide sequence is a sense polynucleotide, an antisense polynucleotide or a ribozyme.
 - 10. A DNA molecule according to claim 7 comprising a polynucleotide sequence encoding a polypeptide capable of disrupting the DNA mismatch repair system of a plant.

- 11. A DNA molecule according to claim 10 wherein said polypeptide is homologous to AtMSH3 (SEQ ID NO: 19) or to AtMSH6 (SEQ ID NO: 31).
- 12. A DNA molecule according to claim 10 further comprising a regulation element capable of causing overexpression of said polypeptide in a cell of said plant.
- 13. A chimeric gene comprising:

5

a DNA sequence selected from the group consisting of (i) a sequence encoding a polynucleotide capable of interfering with the expression of a plant polynucleotide sequence encoding a polypeptide which is homologous to a mismatch repair polypeptide of a yeast or of a human and thereby disabling said plant polynucleotide sequence, and (ii) a sequence encoding a polypeptide capable of disrupting the DNA mismatch repair system of a plant; and

at least one regulation element capable of functioning in a plant cell.

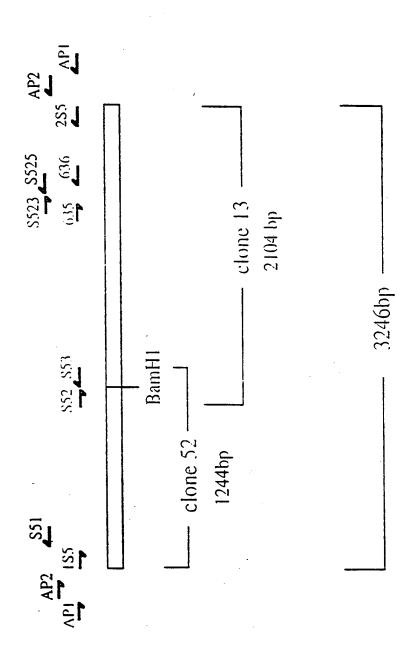
- 14. A chimeric gene according to claim 13 wherein said regulation element is selected from constitutive, inducible, tissue type specific and cell type specific promoters.
- 15. A chimeric gene according to claim 13 comprising a DNA sequence encoding a polypeptide capable of disrupting the DNA mismatch repair system of a plant, wherein said regulation element is capable of causing overexpression of said polypeptide in a cell of said plant.
- 16. A chimeric gene according to claim 13 wherein said regulation element is selected from the group consisting of 35S, NOS, PR1a, AoPR1 and DMC1.
 - 17. A plasmid or vector comprising a chimeric gene according to any one of claims 13-16.
 - 18. A plant cell stably transformed, transfected or electroporated with a plasmid or vector according to claim 17.
- 25 19. A plant comprising a cell according to claim 18.
 - 20. A plant according to claim 19 selected from plants of the families Brassicaceae, Poaceae, Solanaceae, Asteraceae, Malvaceae, Fabaceae, Linaceae, Canabinaceae, Dauaceae and Cucurbitaceae.
- 21. A process for at least partially inactivating a DNA mismatch repair system of a plant cell, comprising transforming or transfecting said plant cell with a DNA molecule according to any one of claims 1-3 or 7-12 and causing said DNA sequence to express said polynucleotide or said polypeptide.
 - 22. A process for at least partially inactivating a DNA mismatch repair system of a plant cell, comprising transforming or transfecting said plant cell with a chimeric gene

according to any one of claims 13-16 and causing said DNA sequence to express said polynucleotide or said polypeptide.

- 23. A process for at least partially inactivating a DNA mismatch repair system of a plant cell, comprising transforming or transfecting said plant cell with a plasmid or vector according to claim 17 and causing said DNA sequence to express said polynucleotide or said polypeptide.
- 24. A process for increasing genetic variation in a plant comprising obtaining a hybrid plant from a first plant and a second plant, or cells thereof, said first and second plants being genetically different; altering the mismatch repair system in said hybrid plant; permitting said hybrid plant to self-fertilise and produce offspring plants; and screening said offspring plants for plants in which homeologous recombination has occurred.
- 25. A process according to claim 24 wherein a first gene is incapacitated in said first plant, a second gene is incapacitated in said second plant, and said first and second genes are incapacitated in said hybrid plant thereby altering the mismatch repair system of said hybrid plant.
 - 25. A process according to claim 25 wherein said incapacitation of the mismatch repair system of said hybrid plant is reversible.
- 26. A process according to claim 24 wherein a new genetic linkage of a desired characteristic trait or of a gene which contributes to a desired characteristic trait is 20 observable in at least one of said offspring plants.
- 27. A process for obtaining a plant having a desired characteristic, comprising altering the mismatch repair system in a plant, cell or plurality of cells of a plant which does not have said desired characteristic, permitting mutations to persist in said cells to produce mutated plant cells, deriving plants from said mutated plant cells, and screening said plants for a plant having said desired characteristic.
- 28. A process according to claim 27 wherein said step of altering the mismatch repair system comprises introducing into said hybrid plant, plant, cell or cells a chimeric gene according to claim 13 and permitting the chimeric gene to express a polynucleotide which is capable of interfering with the expression of a plant polynucleotide sequence in a mismatch repair gene of the hybrid plant, plant, cell or cells, or a polypeptide capable of disrupting the DNA mismatch repair system of the hybrid plant, cell or cells.
 - 29. A process according to claim 28 comprising inactivating an MSH3 gene and/or an MSH6 gene of said plant.
- 30. A process according to claim 28 comprising inactivating an MSH3 gene and an MSH6 gene of said plant.

- 31. A process according to claim 27 comprising at least partially inactivating the mismatch repair system of said plant in a predetermined cell type or in a predetermined tissue of said plant.
- 32. A process according to claim 31 further comprising restoring mismatch repair 5 in said cell type or said tissue.
 - 33. An oligonucleotide capable of hybridising at 45°C under standard PCR conditions to a DNA molecule according to claim 1 with the proviso that said oligonucleotide is other than SEQ ID NO:1 or SEQ ID NO:2.
- 34. An oligonucleotide capable of hybridising at 45°C under standard PCR conditions to the DNA of SEQ ID NO: 18 with the proviso that said oligonucleotide is other than SEQ ID NO:1 or SEQ ID NO:2.
 - 35. An oligonucleotide capable of hybridising at 45°C under standard PCR conditions to the DNA of SEQ ID NO:30 with the proviso that said oligonucleotide is other than SEQ ID NO:1 or SEQ ID NO:2.

Figure 1



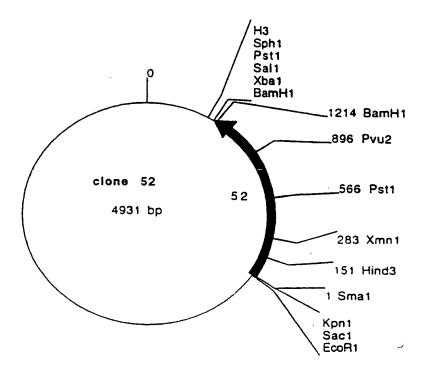


Figure 2

Comments/References: 52= 3' side of S5 (AtMSH3) 1244bp in pUC18/Sma1

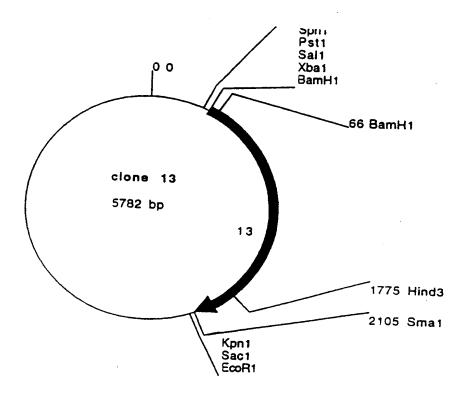


Figure 3

C mments/References: 13 = 3' side of S5 (AtMSH3) 2104bp in pUC18/Sma1

-
0
Ħ
2
$\boldsymbol{\sigma}$
٠H
(Tax

80	144	204 35	264 55	324 75	384 95	444 115	504 135	564 155	624 175	684 195	744 215	804 235	864 255	924 275
AGCGCGCGAAAATTGGCAACCCAAGTTCGCCATAGCCACGACCACGACCTTCCATTTCTCTTTAAACGGAGGA	CCC P	000 P	CTC	GAT D	GTT V	AAG K	GGA G	TTC F	GCA A	GCA A	GAA E	AAT N	GAA E	GTT V
ACG	GCT A	CCA P	CAC H	S P	TAT Y	CTA L	TTC F	NAT N	AAT N	GGT G	CTT L	AGT S	ATT I	GTT V
TTAA	TTC F	000 000	GAC D	GTA V	GAA	GAG E	TTC	CAC H	GTG V	CAT H	ACG T	CAG Q	GGT G	GAA E
TCTC	TTC F	ACA T	TCC	CCA P	GAG E	GTG V	AGA R	GAT D	CTG L	TCC S	200 8	TCA S	TGT C	GGT
CATT	CGT R	TCA S	CTC L	AAC N	SCG P	GTG V	TAC Y	ATG M	AGA R	AAG K	AAA K	GGT	၁၅၅	ACA T
crrc	TCT	TCA S	CTT L	CAA O	TCG S	CAA	AGG R	CAT H	AGA R	ATT I	ACC T	TTT F	TTA L	TCG S
CGAC	ATT I	GAA E	AAG K	ACT	0 0 0	CAG Q	TAC Y	GCT A	GTG V	GCC A	TAT Y	GGT G	ACA T	AT'T I
ACCA	ACG	9 8 9	cgt R	CAC H	GAA E	GZA	GGT G	TAC Y	CAT H	GCA A	TTG L	GAA E	GAG E	GAA E
CACG	CAG Q	GT'A V	AAG K	CCT P	CTG L	TTG L	GTT V	AT'F I	TTC F	ACT T	GCG A	GAA E	TCG S	GTT V
TAGC	CAG Q	000 6	TCC S	TCT S	TTT	CCA P	GAA E	GGT G	AAT N	GAA E	TCG S	GGT G	AAG K	ე ე
CCCA	A.A.G.	AAT N	CCT	CTT L	AGA R	ACA T	GTG V	TTG L	TTG L	ACT T	CTG I.	GGT	GrT V	GTT V
GTTC	CAA	 	TCT	AAA K	CAG 0	TAC Y	ATG M	GTG V	CGA R	CAG Q	GCA G	TGT C	AGA R	GTT V
CCAA	A.A.G K	GAA E	TTC F	CCT P	CTC L	A.A.A	TTG L	CGC R	TTT F	AAG K	CGG R	GGT G	GAG E	GGT G
CAAC	၁၅၅	CAC H	TCC S	AAG K	TTT E	AGG R	GTT V	GCA A	ACA T	GTG V	TTC F	GGT G	GAT D	GTC V
TTGG	ATG M	ACT T	GTA V	A AA	AGA R	TCG S	GTG V	GCA A	CCA P	GTA V	TTT F	AGT S	GTG V	AGA R
AAA	WATT	000 8	ACT T	000 4	c A	TCA S	GAT D	ATC I	GTG V	GGT G	CCT P	ATA	GTT V	GTT V
ງວງວ	ATAAAGCAATT	TCC S	9 8 8	TCA S	CAC H	TCA S	CCA P	GAG E	AGT S	ATT I	၁၅၅	GAT D	TGT C	GAT D
AGCG	ATAA	X X	TCC	GCG A	TTA L	ACG T	TAC	900 A	GCG A	A.A.G K	ACC T	GAG E	GTT V	TTT F
AGAA	•))	ATA I	000 8	N N	GAA E	AAG K	GAC	ACG T	TAC Y	CGG ₽	GCT A	TTG L	AGT S
CCTAAGAA	GATTACG	A AA	AAG K	225	200 d	000 6	AGC S	GAA	ATG M	GGA G	AAC N	GCG	TTC F	ATG M
-	81 1	145 16	205 36	265 56	325 76	· 385 96	445 116	505 136	565 156	625 176	665 196	745	805 236	865 256

GAA GAG TTC AAT GAT AAT TTC ATG E F N D N F M CCA GCT GAG CTG TTG CTT GGC CAG P L L L G Q CAT GCT GGA CCT ACC TCA AAC GTT M A G P T S N V	TTC AAT GAT AAT TTC GAG CTG TTG CTT GGC E L L L G GGA CCT ACC TCA AAC G P T S N	AAT GAT AAT TTC CTG TTG CTT GGC L L L G CCT ACC TCA AAC P T S N	N F CTT GGC L G TCA AAC S N	TTC F GGC G		ATG M CAG O O O		AGA R CCT P CGA R	AGT S CTT L	GGA G TCA S GAA E	TTA L CAA Q CGT R	GAG CAA O O GCC	GCT A ACT T TCA S	GTG V CAG GAG E	ATT I AAG K K GAT D	TTG L TTT F C	AGC S TTG L TTC F	TTG L GTG V AGC S	984 295 1044 315 1104 335
5 5		A A A A	GTA V GAA E	GAT D ATG M	GAG E AAG K						GAA E AAA K	AAA K GGA G	ATC I ATG M	AGC S TCT S	GCA A TGC C	GGT G TTG L	AAC N ACA T	TTA L GTT V	1164 355 1224 375
	ATT I CAG	ATG M TTT	AAC N GGA G	ATG M TTT	CCA P GAA E	CAT H AGG R	CTG L ATC I	ACT T CTT L	GTT V V TAC	e de de de	0000 0000 0000	CTC L GCC	GCC A TCA S	CTA L TTT F	ACG T CGC	TTT F TCT S	TGC C TTG	CAT H TCA S	1284 395 1344 415
AAC P	ACA T GAT	GAG E GGA	ATG M TCG	ACT T GAA	CTC	TCA S GGC	GCC A TCC	AAT N TTA	ACT T TTC	CTG L CAT	CAA O AAT	CAG Q ATG	TTG L AAT	GAG E CAC	GTT V ACA	GTG V CTT	AAA K ACA	AAT N GTA	1404 435 1464
-	rcc s		CTT L							н САТ Н	r CCT						T TTG L	v ATA I	455 1524 475
	R R AGC	CTT L AGT	GAT D GAG	GCT A TTG	GTT GTT	TCT S GAA	GAG GAA	ATT I GGT	rcr S rcr	GCT A GAG	TGC C AGA	M M GCA	GGA G ATT	TCT S GTA	CAT H TCA	S S CCT	rcr s gag	S S TTT	1584 495 1644
	s GTG V																cgr R	GGA G	1704 535
ACA 1	AGA R	ATC I	TTT F	CAT H	000 8	ACT	CCT	AAA K	225 4	ACA	GAG	TTC	ATT	GCA A	GTT V	ATG M	GAA E	GCT A	1764 555
TTA (TI	900 4	999	AAG K	CAA	ATT	CAG Q	s cgc crr R L Figure	Crr L	GGC ATA AAG CAA G I K 0 4 (Continued)	ATA I onti	AAG K •	CAA G	GAC	S	GAA	ATG M	AGG R	1824 575

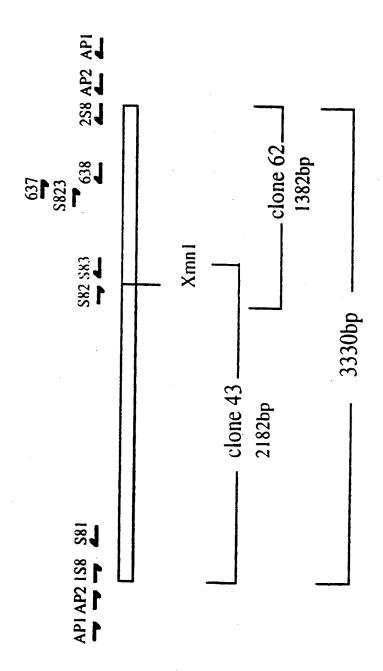
1825 576	AGT S	ATG M	CAA	TCT S	GCA	ACT T	GTG V	CGA 7	TCT	ACT (CTT	TTG /	AGA 1	AAA 7	TTG #	ATT 1 I	TCT (S	GTT A	ATT I	TCA S	1884 595
1885 596	TCC S	CCT P	GTT V	GTG V	GTT V	GAC D	AAT N	9 P	GGA	AAA K	CTT (CTC 1	TCT (S	GCC (CTA I	AAT I	AAG (K	GAA (900 Y	GCT A	1944 615
1945 616	GTT V	CGA	GGT G	GAC D	TTG	CTC	GAC D	ATA (CTA	A'I'C	ACT T	TCC 1	AGC (S	GAC (CAA 1	TTT (F	CCT (GAG (E	CTT (GCT A	2004 635
2005	GAA E	GCT A	င်င်င န	o caa	GCA A	GTT V	TTA L	GTC .	ATC	AGG (GAA :	AAG (CTG (GAT O	TCC 1	rcc 1	ATA (I	GCT 7	TCA	TTT E	2064 655
2065 656	000 8	A AG	A AG	CTC L	GCT	ATT I	CGA R	AAT	TTG L	GA.A	TTT (CTT	CAA	GTG 7	TCG (S	999	ATC I	ACA (CAT	TTG L	2124 675
2125 676	ATA I	GAG E	CTG L	000 a	GTT V	GAT D	TCC S	AAGK	GTC V	CCT P	ATG	A.Y.	TGG W	GTG ,	AAA (GTA V	AAT	AGC S	ACC	A.A.G K	2184 695
2185 696	AAG K	ACT T	ATT I	CGA R	TAT Y	CAT H	႕ ၁၁၁	CC.A	GAA E	ATA I	GTA V	CCT.	 	91.4 1	GAT (GAG	CTA (GCT	CTA	GCA A	2244 715
2245	ACT	GAA E	CAT H	CTT L	OCC P	ATT I	GTG V	AAC N	CGA R	GCT A	TCG S	TGG W	GAT D	AGT	TTC (CTC	AAG K	AGT	TTC	AGT S	2304 735
2305 736	AGA R	TAC Y	TAC Y	ACA T	GAT D	rtt F	AAG K	GCT A	OCC V	GTT V	CAA	GCT A	CTT L	GCT	GCA	CTG	GAC	TGT	TTG L	CAC H	2364 755
2365 756	TCC S	CTT L		ACT T	CTA L	TCT S	AGA R	AAC	AAG K	AAC N	TAT Y	GTC V	CGT.	222	GAG '	TTT	GTG V	GAT	GAC	TGT C	2424 775
2425	GAA E	CCA P		GAG E	ATA I	AAC N	ATA I	CAG Q	TCT S	GGT G	CGT R	CAT H	CCT P	GTA V	CTG (GAG	ACT	ATA	TTA	CAA Q	2484 795
2485 796	GAT D	N A		GTC V	CCA P	AAT N	GAC D	ACA T	ATT I	TTG L	CAT H	GCA A	GAA E	9 999	GAA	TAT	TGC C	CAR	ATT	ATC I	2544 815
2545 816	ACC	GGA G	CCT	AAC N	ATG M	GGA G	GGA G	A.A.G.	AGC S	TGC C	TAT Y	ATC I	CGT R	CAR	GTT V	GCT A	TTA L	ATT	TCC	ATA I	2604 835
2605 836	ATG M	GCT A	CAG Q	GTT V	GGT G	700 8	TTT F	GTA V	CCA P	9	TCA	TTC F	9 8	AAG K	CTG	CAC	GTG V	CTT	GAT	GGT	2664 855
2665 856	GTT V	TTC	ACT	000 8	ATG	GGT G	GCT A	TCA	GAC D	S AGT A'S I	ATC I	CAG	CAT H	GGC AGA Ge R :inued)	_	AGT	ACC T	TTT F	CTA L	GAA	2724 875

2725 876	GAA	TTA L	AGT	GAA	GCG A	TCA S	CAC	ATA	ATC	AGA R	ACC T	TGT 1	rcr 1	TCT (S	CGT 1	TCG (S	CTT (GTT /	ATA	TTA L	2784 895
785 896	GAT D	GAG E	CTT L	GGA G	AGA R	၁၅၅	ACT	AGC	ACA	CAC	GAC (GGT (GTA (gcc 1	ATT (I	GCC 1	TAT (Y	GCA J	ACA	TTA L	2844 915
916	CAG Q	CAT H	CTC L	CTA	GCA A	GAA E	AAG K	AGA R	TGT	TTG	GTT (CTT 1 L	TTT (GTC 1	ACG (CAT 1	TAC (Y	CCT (GA.A	ATA I	2904 935
2905 936	GCT A	GAG E	ATC I	AGT S	AAC N	GGA G	TTC F	CCA	GGT	TCT	GTT (000	ACA T	TAC (CAT (GTC 1	TCG 1	TAT (Y	CTG	ACA T	2964 955
965 956	TTG L	CAG	AAG K	GAT D	A A	၁၅ ၁	AGT S	TAT Y	GAT D	CAT	GAT	GAT (GTG /	ACC 1	TAC	CTA 7	TAT /	AAG (CTT L	GTG V	3024 975
1025 976	CGT R	GGT	CTT L	76C C	AGC S	AGG R	AGC S	TTT F	GGT G	TTT F	AAG K	GTT (GCT (CAG	CTT) v	CAG 1	ATA (CCT P	CCA	3084 995
996 996	TCA	TGT C	ATA I	CGT R	CGA R	သ ဗ	ATT I	TCA	ATG M	GCT.	GCA A	AAA	TTG (GAA (GCT (GAG (GTA (CGT (GCA A	AGA R	3144
1145	GAG E	AGA R	AAT	ACA T	င် (၁	ATG M	GGA G	GAA E	CCA P	GA.A	GGA	CAT (GAA (GAA C	CCG P	AGA (, 5 5 5 5 5 5 5	GCA O	GAA E	GAA E	3204 1035
1205 .036	TCT	ATT I	TCG S	GCT A	CTA L	GGT	GAC	TTG L	TTT E	GCA A	GAC	CTG /	X X	TTT (GCT A	CTC 3	TCT (S	GAA (GAG E	GAC	3264 1055
3265 1056	CCT P	TGG ¥	X X	GCA A	TTC F	GAG E	TTT F	TTA L	AAG K	CAT	GCT	TGG 1	AAG K	ATT (GCT (2000	AAA /	ATC 1	AGA R	CTA L	3324 1075
1325	A A	CCA P	ACT	TGT	TCA S	TTT F		TTTA	MTCT	TAAC	ATTA	TGA TTTAATCTTAACATTATAGCAACTGCAAGGTCTTGATCATCTGTTAGTTGCG	A.A.C.T.	GCNA	GGTC	TTGA'	rcati	crgr.	TAGT	TGCG	3397 1082
1398 1	TAC.	TACTAACT	_	ATG TC	TGT AT C I	ATT AG	AGT AT S I	ATA AC I T	ACA AGA AAA T R K	. A.A. K	A GA	GAG AAT 1	T T.A.	G AG	TAG AGAG ATG	TG G	GAT TO	TCT A	AAT C	900 a.	3458 5
3459 6	org V	TTG L	CAG	TAC Y	ATC	TTT	S	CCA P	ည သသ	C GCA TAA A Figure	TAA A		qaaaaaaaaaaa (Continued)	AAAA .nue	AAAA d)	AAAA	AAAA	AAAA	\$		3522 16

c amig

1. At 1 MEKOK	1 At 52 BDELARASPRESTORS 3 GC 96 VOURNERMENTSTRESTORS	AC 134 FORDARIA IN PUBLICA 1. So 191 EARDAVY SELEMENT	3_AC 213 MLEAARDISGGCGGEEGFGGGSMFLVCVVDRRVKSKTGGCGIEMSFUVWOVVOVVOVETSETSTSTVV 3_8C 206 NFOVMSTFVLRGKR	3_AK 306 BSQCTEKTLVAHAOPTSMVRVERASLDCYSNGNAVDBVKSBCCKISACNLEDDKKNKLKAARKGMSCLTVBTIMNMPBLTVQALALTFCER 3_kc 368 BPLBVAKSPKDISCPLISKCKYDLEDHVVCAIKVMNKKBOBSPEL····································) AC 401 PENTENDOASUSESS.	NAC 696 OLBSRLVKROSKRAIV Nac 517	3_Ac 590 (100)	LAC 600 VDSKVDWNZVSVEST	3 At 775 CE	H) AC 870 RSTRUKELBERASHITATUSSRSTVIDDELGROTSTEDOVLIA VATLOHLIAEKRO-TVETVETTVOKIA DIBNOFRSVOTVIDAKUR OKLONOM H) sc. 873 DSTERVOHLDILHIKRAMAKRELLIDEVOROTUTHOOIAIS VALLAVESKLEDER KETTHVOHLOSIKSPLIMMKHMONVEKOK	2 2	
1 1	2 x 3	8 E 3		1.1		1.1		8 E 3	2 2 2 3 3 3	S 23	~ ~	•

Figure 6



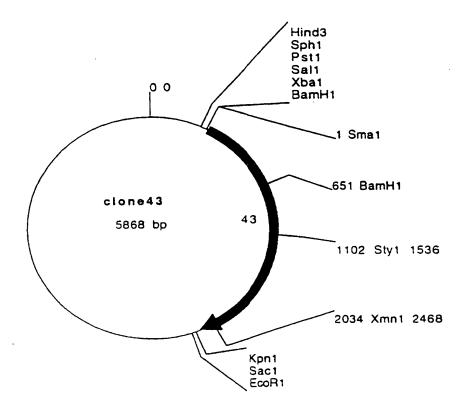


Figure 7

Comments/References: 43= 5' side of S8 (AtMSH6) 2182 bp in pUC18/Sma1

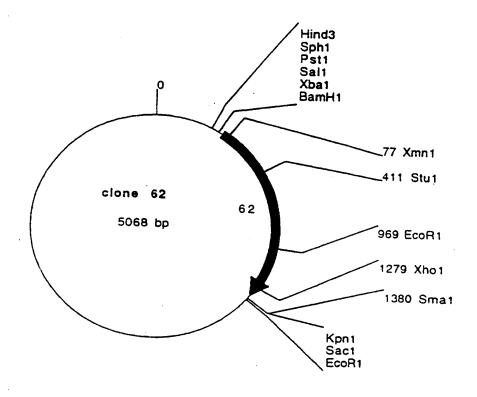


Figure 8

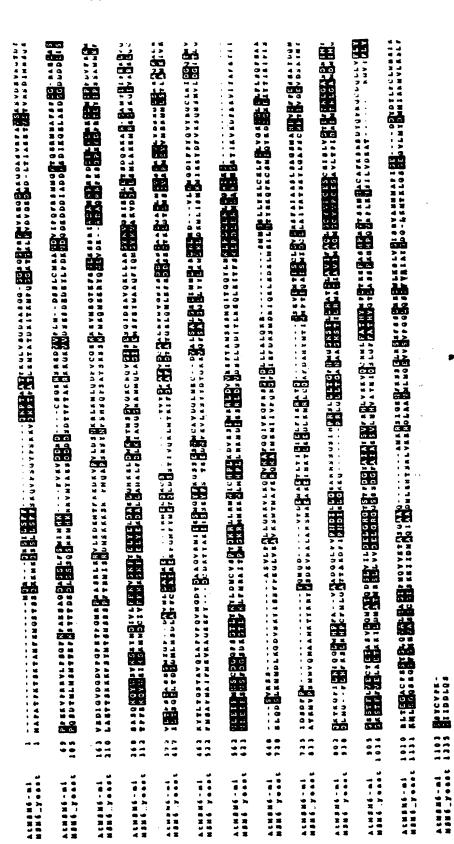
Comments/References: 62= 3' side of S8 (AtMSH6) 1379bp in pUC18/Sma1

TCTCTCAGCTCAAAACATCGTTTCTCTCTCTCTCTCTCAAAATTCCAAAAA ATG CAG CGC CAG 153
G CT TTG GTT TCC
TTG L AAG R
ACG T TTT F TCT
GCG A CCA P
ACC GCG T A GGA GGA
.
CAA AAA
TTC TTC
ר ביטני טניני
TTTCGAAT
81 TJ

AAG AAA ATG TCT K K M S	۲	E	ງໄດ້ ຮ	_	SCA A	a .	()	_	=				a .	C)	_	•	(')	GAC 1	ATT I	993 284
GTG CTT TTC TTT AAA GTG GGG AAA V L F F K V G K	TTT AAA GTG GGG F K V G	TTT AAA GTG GGG F K V G	TTT AAA GTG GGG F K V G	AAA GTG GGG K V G	999 9	₹ ⊁		TTT F	TAT Y	GAG E	CTG L	TAT Y	GAG E	CTA L	GAT	GCG A	GAA E	TTA L	GGT G	1053 304
CAC AAG GAG CTT GAC TGG AAG ATG	CTT GAC TGG AAG L D W K	CTT GAC TGG AAG L D W K	CTT GAC TGG AAG L D W K	TGG AAG W K	AAG K	Σ		ACC T	ATG	AGT S	ggr G	GTG V	GGA G	AAA K	TGC C	AGA R	CAG Q	GTT V	GGT G	1113 324
AGT GGG ATA GAT S G I D	AGT GGG ATA GAT S G I D	AGT GGG ATA GAT S G I D	AGT GGG ATA GAT S G I D	ATA GAT I D	GAT D	© ш	GAG	GCA A	GTG	C G	AAG K	C'FA L	TTA L	GCT A	CGT .	GGA G	TAT Y	A A	GTT V	1173
GAG CAG CTA GAA E Q L E	GAG CAG CTA GAA E Q L E	GAG CAG CTA GAA E Q L E	GAG CAG CTA GAA E Q L E	CTA GAA L E	GAA E	ĕ ⊢	ACA			O CAN			GCC A	AGA R	GGT	GCT A	AAT	ACT T	ATA I	1233 364
AAG CTA GTT CAG K L V Q	AAG CTA GTT CAG K L V Q	AAG CTA GTT CAG K L V Q	AAG CTA GTT CAG K L V Q	CTA GTT CAG L V Q	CAG Q	$\sigma >$	GTA '	TTA L	ACT	CCA P	TCA S	ACA T	GCA A	AGC S	GAG (GGA G	AAC N	ATC I	999 999	1293 384
GTC CAT CTT CTT	GTC CAT CTT CTT	GTC CAT CTT CTT	GTC CAT CTT CTT	CAT CTT CTT H L L	CTT	ত ৰ	GCT ,	ATA I	AAA	GAG	ATC I	X Z	ATG M	GAG E	CTA (CAA	AAG K	TGT C	TCA S	1353 404
ACT GTG TAT GGA TTT GCT TTT GTT T V Y G F A F V	GGA TTT GCT TTT G F A F	GGA TTT GCT TTT G F A F	GGA TTT GCT TTT G F A F	TTT GCT TTT F A F	TTT F	ច >		GAC	TGT	GCT	000 A	TTG	AGG R	TTT F	TGG W	GTT (999 9	TCC	ATC	1413 424
AGC GAT GAT GCA TCA TGT GCT GCT S D D A S C A A	GCA TCA TGT GCT A S C A	GCA TCA TGT GCT A S C A	GCA TCA TGT GCT A S C A	TCA TGT GCT S C A	GCT A	७ द		CTT	GGA G	gcc A	TTA	TTG	ATG	CAG Q	GTT V	TCT (CCA	AAG K	GAA	1473 444
GTG TTA TAT GAC AGT AAA GGG CTA V L Y D S K G L	GAC AGT AAA GGG D .S K G	GAC AGT AAA GGG D .S K G	GAC AGT AAA GGG D .S K G	AAA GGG K G	ე ე	Σ \rightarrow		TCA	AGA (GAA	GCA O	CAA	AAG K	GCT.	CTA /	AGG 1	AAA	TAT Y	ACG T	1533 464
TTG ACA GGG TCT ACG GCG GTA CAG	rcr Acg GCG GTA S T A V	rcr Acg GCG GTA S T A V	rcr Acg GCG GTA S T A V	GCG GTA	GTA V	ပ်ထ		TTG (GCT (CCA (GTA (CCA (CAN	GTA V	ATG (9 999	GAT	ACA T	GAT D	1593 484
STT AGA AAT ATA	STT AGA AAT ATA	STT AGA AAT ATA	STT AGA AAT ATA	AAT ATA N I	ATA I	¥Η	ATA (I	GAA :	TCT /	AAC (GGA '	TAC	TTT /	A A A	GGT 1	rcr 1 S S	TCT (GAA E	TCA S	1653 504
TGG AAC TGT GCT GTT GAT GGT CTA	SCT GTT GAT GGT	SCT GTT GAT GGT	SCT GTT GAT GGT	GAT GGT D G	GGT G	Cl		AAT (GAA 1	TGT (GAT (GTT (9 225 GCC	CTT 1	AGT G	GCT C	CTT (GGA	GAG E	1713 524
CTA ATT AAT CAT CTG TCT AGG CTA	CAT CTG TCT AGG	CAT CTG TCT AGG	CAT CTG TCT AGG	TCT AGG	AGG R	CŢ		AAG C K	CTA C	GAA (GAT (GTA (CTT A	AAG C	CAT G	0 0 0 0 0 0 0	GAT /	ATT	TTT F	1773 544
CCA TAC CAA GTT TAC AGG GGT TGT P Y Q V Y R G C	STT TAC AGG GGT	STT TAC AGG GGT	STT TAC AGG GGT	AGG GGT R G	GGT G	၌		CTC A	AGA ATT R I Figure		GAT (D) (S)	AT GGC CAG P G Q T (Continued)	CAG P	90	ATG G M V	GTA A V	AAT O	CTT	GAG E	1833 564

1893 584	1953 604	2013 624	2073 644	2133 664	2193 684	2253 704	2313 724	2373 744	2433 764	2493 784	2553 804	2613 824	2673 844	2733 864
A.A.C.	GAT D	AGT S	CGC R	GTG V	GAT D	A AAA	GCC A	ACT T	CAC H	GGA G	ACA T	GGT G	ATT I	CGT R
GAT D	A A A	GAA E	GGA G	A A	ATT I	TGT C	GCA A	GAA E	ATT	GCT A	AAA K	GAT D	AGC S	CTT L
CTT	CTC	TCA S	CTC L	¥ ¥	GGA G	стс L	GAA E	GCT A	GTC V	TCT S	CAG Q	9 8 9	၁၅၅	CTT L
TAT	CCA P	AAC N	CTG L	9 9 9	AGT S	ANA K	TTC F	AAC N	GAG E	CTC L	AAT N	GCA A	AGT S	ACT T
AAA K	CAT H	GCA	AGA R	CTG L	AGA R	TAT Y	CAA	GAA	TCT S	AGT S	CAG Q	GTT V	AGC S	TCA
TAC	TGC	ACG T	GAA	Crt	1.1C F	CTT L	TCT S	GAT D	TGG ₩	GCA A	GAT D	GCA A	AGA R	AAA K
TTG . L	ATC I	TTC	T'fA L	GCT A	999 9	17FG L	CTT [.	ACA T	O CAN	GCA	ACA T	TTT F	AGA R	GGA OP nued
ACC T	TGG W	GAA E	GAC D	CC:T P	A A	S S	rrr F	GTG V	ACT T	ATC I	GCT A	CCA P	GCT A	16 GGC G (Conti
999	AAT.	GAA E	CCA P	TTG L	GTC V	ATG M	TTA L	GAT D	GCA A	GCA A	GAA	CAT H	GAG E	ATG M
TCA	AGG R	GTT V	CTT L	GTG V	ATT I	ATG M	GAG E	CAA	AGA R	TTT F	TCA S	TGG ₩	၁၅၁	AAC, N Nre
CCT	T'FA L	GTA V	¥ X	TCT	CAA	AA'I' N	C'FA L	AAC	GAA E	TCT S	GAA E	CTA L	CTT L	CCA P Fig
GGT G	CTC L	GAT. D	CAC H	OCC A	000 0	TCA S	999	CAG 0	ATC 1	AGA R		GGA G	CTC L	GGA G
GGT (CGA R	CTT L	CTC 1.	TCA S	ፒነገ F	GAA	AGC S	TAT Y	TTT F	CTG L	TTT F	CAA O	ATA I	ACG T
GAT O	AAG K	CGG R	TAT Y	TCA S	GCA A	AAG K	A A A A	AAT N	CTT L	GTC V	ATT I	ATC I	GAT D	CTG L
TGT (GGT .	X AA	CAG Q	CGA R	X X X	CAG 0	GGA G	CCA	GAA E	GAT D	GTT V	AAA K	AAT N	TTA L
AGC '	ACT	AAT	၁၁၅	GTT V	GTT V	CTA L	GTA V	TTT F	ATC I	CTA L	CCT P	CTT L	႕ ၅၁၁	TTG L
AAT	CCA .	ATC	ACT	AGC S	CGA R	GCT A	TTA L	GAC	CTT	16C C	AGG R	ATA I	GTT V	TCA
AAC A	AGT	AGC S	ATC	TCT	CAA	TTG L	ATA	AGC S	ATA I	AGC S	000 8	CCA P	CCT	992 %
TTT F			CAA					GAT D		ATA I	ATG M	999 9	r TTG L	CCT
ATA 1 I	TGT	GTA (CTG			CTC	ACC	AGC	A A A	CAA O	CAT H
1834 <i>1</i> 565	1894 7					2194			2374	2434 765	2494	2554 805	2614 825	2674 845

2734 865	GCA A	ACA T	TGT	CTG L	gcc •	GTT V	ATC I	TTT F	000 A	CAA	CTT	၁ဗ္ဗ	76C	TAC Y	GTG V	, 500 P	TGT	GAG	TCT S	76C C	2793 884
2794 885	GAA E	ATC I	TCC	CTC L	GTG V	GAT D	ACT T	ATC I	TTC F	ACA	AGG R	CTT	299	GCA A	TCT	GAT	AGA R	ATC	ATG	ACA T	2853 904
2854 905	GGA G	GAG E	AGT	ACC	TTT F	TTG L	GTA V	GAA E	1GC C	ACT T	GAG E	ACA	ece A	TCA S	GTT V	CTT	CAG	AAT	GCA	ACT T	2913 924
2914 925	CAG	GAT D	TCA S	CTA	GTA V	ATC I	CTT L	GAC D	GAA E	CTG L	၁၁၁	AGA R	GGA	ACT	AGT	ACT	TTC	GAT	GGA G	TAC Y	2973 944
2974 945	005 4	ATT I	GCA A	TAC Y	TCG S	GTT	TTT F	cgt R	CAC H	CTG L	GTA V	GAG E	AAA K	GTT V	CAA	TGT	CGG	ATG	CTC	TTT	3033 964
3034 965	GCA A	ACA T	CAT H	TAC Y	CAC H	CCT P	CTC L	ACC T	AAG K	GAA E	TTC F	GCG A	TCT S	CAC	CCA P	CGT R	GTC V	ACC	TCG	AAA K	3093 984
3094 985	CAC	ATG M	GCT A	7GC 0	GCA A	TTC	A AA	TCA S	AGA R	TCT S	GAT D	TAT Y	C AA	CCA P	CGT R	GGT	TGT	GAT	CAA	GAC D	3153 1004
3154 1005	CTA L	GTG V	TTC F	TTG L	TAC Y	CGT R	TTA L	ACC T	GAG E	GGA G	GCT A	TGT C	CCT P	GAG E	AGC S	TAC	GGA	CTT	CAA	GTG V	3213 1024
3214 1025	GCA	CTC L	ATG	GCT A	GGA G	ATA I	CCA P	AAC N	CAA O	GTG V	GT'ſ V	GAA E	ACA	GCA A	TCA S	GGT	GCT	GCT A	CAA	9 ₽	3273 1044
3274 1045	ATG M	AAG K	AGA R	TCA S	A'I''ſ I	ე ეეე	GGA E	AAC N	TTC F	AAG K	TCA	AGT S	CAG	CTA L	ACA R	TCT	GAG	TTC	TCA	AGT S	3333 1064
3334 1065	CTG L	CAT H	GAA	GAC	TGG W	CTC L	AAG K	TCA S	TTG L	GTG V	GGT G	ATT I	TCT	CGA R	GTC V	GCC A	CAC	AAC	AAT	200 A	3393 1084
3394 1085	CCC P	ATT I	၁၅၅	GAA	GAT D	GAC	TAC Y	GAC D	ACT T	TTG L	TTT F	7GC C	TTA	TGG W	CAT	GAG .	ATC	AAA K	TCC	TCT	3453 1104
3454 1105	TAC Y	TGT C	GTT V	<u>ရ</u>	A X	TAA	ATG M	GCT.	ATG M	ACA T	TAA		ATCT	GAAG	CTCG	TTAA	GTCT	CACTATCTGAAGCTCGTTAAGTCTTTTGCCTCTCT	CCTC	TCT	3521 5
3522 1	ά Σ	ATG T'	TTT A' F I	ATT CO	CCT CI	CTT A	AAA AA K	AAA TO	TGC TT C L	TTA TAT L Y		ATC AA I K	AAA AAA K K	A TTG L	G TTT F	T CCT	T CGA R	A TTA L	A AAA K	æ	3579 19
3580 20	AAA K	A AA	K AA	K AA	K K	K AA	AAA K	AAA K	AAA K	-	Figure	<u>ა</u>	(Continued)	penu	_						3606 28



·	
TTTTTTGGTTGCTAACAATAAAGGTATACGGTTTTATGTCATCAATATAA	50
CTATATATAAAAGAAATGAAAGATATATATTGTTTTTTCATTTATCAAAC	100
AAAACAACAAGACTTTTTTTTTTTTTTTTTTTTTTTTTT	150
GATAAACGACATCGTTTAATCATTTCCCAATTTTACCCCTAAGTTTAACA	200
CCTAGAACCTTCTCCATCTTCGCAAGCACAGCCTGATTAGGAACAGCTTT	250
ACCATTCTCATATTCCTGAACTACCTGAGTCCTCTCATTGATCTGTTTCG	300
CCAAATCCGCTTGTGACATCTTCTCCCAATCTCGCTTTCTGTATCATC	350
AACCTCACCTCTGCTTTCACACGATCCATCGCCGCAGGCTCTGTTTCTTC	400
TTCCAGCTTCTTCGTGTTAATCACCGGAACCGCCGTAGATTTCCCCTTTT	450
TGTTCGAACCGGCATCGAATTTCTTAACCGTTTGAACCGCGACACCGTTT	500
CTCAGAGCTGCGTTAACCGCTTTCGGATCGCGTAGGTCTTGGCTCTTTTG	550
TTTTGATTTGTGGAGAACTACTGGTTCCCAGTCTTGTGTTACTGCTCCTG	600
GGTATCTGCTCGGCATCGTCGATGAATTGAGAGAAAGGAACAACGCGAAA	650
ATTTTATTAATCTGAGTTTTGAAATTGAGAAACGATGAAGATGAAGAATG	700
TTGTTGAGAGGATTGTGATATTTATATACGAAGATTGGTTTCTGGAGA	750
ATTCGATCATCTTTTCTCCATTTTCGTCTCTGGAACGTTCTTAGAGATG	800
ATTGACGACGTGTCATTATCTGATTTGCAGTTAACCAATGCTTTTTGGGT	850
TGGATTCGTGGTACACCATATTATCCGATTTGGCTCAATGGTTTTATATA	900
AATTTGGTTTTCGGTTCGGTTATGAGTTATCATTAAAATTAAGCTAACCA	950
AAAATTTTCGTAAAATTTATTTCGGTTTCAATTCGGATCCCTTACTTCCA	1000
GAACCGAATTATTCGAAACCGGGGTTAGCCGAACCGAATACCAATGCCTG	1050
ATTGACTCGTTGGCTAGAAAGATCCAACGGTATACAATAATAGAACATAA	1100
ATCGGACGGTCATCAAAGCCTCAAAGAGTGAACAGTCAACAAAAAAAGTT	1150
GAGCCCTGAGGAGTATCGTTTCCGCCATTTCTACGACGCAAGGCGAAAAT	1200
TTTTGGCGCCAATCTTTCCCCCCTTTCGAATTCTCTCAGCTCAAAACATC	1250
GTTTCTCTCTCACTCTCTCACAATTCCAAAAAATGCAGCGCCAGAGAT	1300
CGATTTTGTCTTTCCAAAAACCCACGGCGGCGACTACGAAGGGTTTG	1350
GTTTCCGGCGATGCTGCTAGCGGCGGGGGGGGGGGGGGG	1400
AATGTGAAGGAAGGGGATGCTAAAGGCGACGCTTCTGTACGTTTTGCTGT	1450
TTCGAAATCTGTCGATGAGGTTAGAGGAACGGATACTCCACCGGAGAAGG	1500
TTCCGCGTCGTGTCCTGCCGTCTGGATTTAAGCCGGCTGAATCCGCCGGT	1550
GATGCTTCGTCCCTGTTCTCCAATATTATGCATAAGTTTGTAAAAGTCGA	1600
TGATCGAGATTGTTCTGGAGAGAGGTACTAATCTTCGATTCTCTTAATTT	1650
TGTTATCTTTAGCTGGAAGAAGAAGATTCGTGTAATTTGTTGTATTCGTT	1700
GGAGAGATTCTGATTACTGCATTGGATCGTTGTTTACAAATTTTCAGGAG	1750
CCGAGAAGATGTTCCCCCTGAATGATTCATCTCTATGTATG	1800
ATGATGTTATTCCTCAATTTCGTTCCAATAATGGTAAAACTCAAGAAAGA	1850
AACCATGCTTTTAGTTTCAGTGGGAGAGCTGAACTTAGATCAGTAGAAGA	1900
TATAGGAGTAGATGGCGATGTTCCTGGTCCAGAAACACCAGGGATGCGTC	1950
CACGTGCTTCTCGCTTGAAGCGAGTTCTGGAGGATGAAATGACTTTTAAG	2000
GAGGATAAGGTTCCTGTATTGGACTCTAACAAAAGGCTGAAAATGCTCCA	2050
GGATCCGGTTTGTGGAGAGAAGAAGAAGTAAACGAAGGAACCAAATTTG	2100
AATGGCTTGAGTCTTCTCGAATCAGGGATGCCAATAGAAGACGTCCTGAT	2150
GATCCCCTTTACGATAGAAAGACCTTACACATACCACCTGATGTTTTCAA	2200

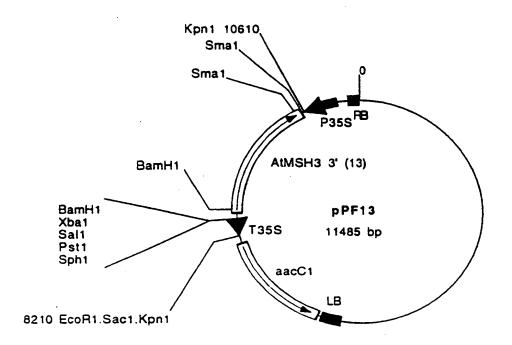
Figure 11

GAAAATGTCTGCATCACAAAAGCAATATTGGAGTGTTAAGAGTGAATATA	2250
TGGACATTGTGCTTTTTTTTTTAAGTGGTTAGTAACTATTAATCTAGTGTT	2300
CAATCCATTTCCTCAATGTGATTTGTTCACTTACATCTGTTTACGTTATG	2350
CTCTTCTCAGGGGAAATTTTATGAGCTGTATGAGCTAGATGCGGAATTAG	2400
GTCACAAGGAGCTTGACTGGAAGATGACCATGAGTGTGTGGGAAAATGC	2450
	2500
AGACAGGTAAATTAGTTGAAACAACTGGCCTGCTTGAATTATTGTGTCTA	2550
TAAATTTTGACACCACCTTTTGTTTCAGGTTGGTATCTCTGAAAGTGGGA	2600
TAGATGAGGCAGTGCAAAAGCTATTAGCTCGTGGGTAAGGGAACCATCAT	2650
ACTITATGGAATTCGTTTACTGCTACTTCGGCTAGGATTTAAGAAATGGA	2700
AATCACTTCAAGCATCATTAGTTAGGATCCTGAGAACTCAGGATGTTTTC	2750
TTATTCGTTATATAATAAGTCTTTTCATCAAGGAGTAACAAACA	2800
GCACAATATTTGTGTGCTCACTGGCAAGGCATATATACCCAGCTAACCTT	2850
TGCTAGTTCACTGTAGTAACAGTTACGGATAATATATGTTTACTTGTATG	2900
TGGTACCCTCATTTTGTCTCTCATGGAGGCTTTCAAGCCTTGTGTTGAAA	2900
CTGGATAGTTACATATGCTTCCAACAGAAACTAGCATGCAGATTCATATG	
CTTTCCTATTCTACTAATTATGTATTGACACACTCGTTGTTTCTTTTGAA	3000
AGATATAAAGTTGGACGAATCGAGCAGCTAGAAACATCTGACCAAGCAAA	3050
AGCCAGAGGTGCTAATACTGTAAGTTTTCTTGGATAGGTCAAGGAGAGTG	3100
TTGCAGACTGTTTTTGATCATTTCTTTTTCTGTACATTACTTTCATGCTG	-3150
TAATTAACTCAATGGCTATTCTGGTCTGATTATCAGATAATTCCAAGGAA	3200
- GCTAGTTCAGGTATTAACTCCATCAACAGCAAGCGAGGGAAACATCGGGC	-3250
CTGATGCCGTCCATCTTCTTGCTATAAAAGAGGTTTGTTATTTACTTATT	3300
TATCTTATCATGTTCAGTTCATCCAAGTCCTGAAAAATTACACTCTTCTT	3350
TACCAATCTTCCATCAAGCTGTGAAAGGATTTGGAATTAGAAAATCATT	3400
ATTTGATGCTTTGTTTTATATGCAAGAGGTTCCCTTGAAAAGATCTGTTT	3450
AAGATTCTTTGCACTTGAAAAATTCAATCTTTTTAAGTGAATCCCCTACT	3500
TTCTTACAATGATCATAGTCTGCAATTGCATGTCAAGTAATATCATTCCT	-3550
TGTTACTGCATCCCCCTCTTTCTTAATGACCATTGTCTATGTTGTGTTTTG	3600
TCTCGTGTGCTGGAGAAAATGATAGCTGATCCAAGCTGTACATTATCATG	3650
ATTAAGTAGCTGCTCAGGAATTGCCTTTGGTTACATTGCCTAATGGTTTG	3700
ATGTCAATTTTCTTCTGAATCTTTATTTTAGATCAAAATGGAGCTACAA	3750
AAGTGTTCAACTGTGTATGGATTTGCTTTTGTTGACTGTGCTTGCCTTGAG	3800
GTTTTGGGTTGGGTCCATCAGCGATGATGCATCATGTGCTCCTCTTGGAG	3850
CGTTATTGATGCAGGTAAGCAAGTGTATTCTGTATCTTATGTGTACCATG	3900
TGACTTCCTGTGCATATATTTGGGTTGCAGGAACTAATTCTGAATCACCA	3950
TTTGGTATGTTTTTCCAGGTTTCTCCAAAGGAAGTGTTATATGACAGTA	4000
AAGGTAAACTGCTTGTATCGCCAGTTGTTTTGTTAAACAGAATTTAAGGT	4050
	4100
AAATGACACTGGTTAATTTAAAGTGCATACATGTTGAAATATTGCAGGGC	4150
TATCAAGAGAAGCACAAAAGGCTCTAAGGAAATATACGTTGACAGGTACC	4200
ATTTCAGTAGGCAAGCTAACTGACAATTTAACCGCTCACCGAATGATAGG	4250
TCTCTTAAACATTGCTAATGTAGATGATGTTTATGTTTCAATCTAATAGG	4300
GTCTACGGCGGTACAGTTGGCTCCAGTACCACAAGTAATGGGGGATACAG	4350
ATGCTGCTGGAGTTAGAAATATAATAGAATCTAACGGATACTTTAAAGGT	4400
TCTTCTGAATCATGGAACTGTGCTGTTGATGGTCTAAATGAATG	

TGCCCTTAGTGCTCTTGGAGAGCTAATTAATCATCTGTCTAGGCTAAAGG	4450
TGTGTTGGCTTGTTTAGTTTTTGCTTTTCACAAATTAAGCAAAGGAACTT	4500
TTCATAACTTACAGTTTCTATCTACTTGCAGCTAGAAGATGTACTTAAGC	4550
ATGGGGATATTTTTCCATACCAAGTTTACAGGGGTTGTCTCAGAATTGAT	4600
GGCCAGACGATGGTAAATCTTGAGATATTTAACAATAGCTGTGATGGTGG	4650
TCCTTCAGGCAAGTGCATATTTCTTTTTTGATAACTTCAACTAGAGGGCA	4700
GACATAGAAGGAAAAATTCTAATACTTCGTACGGATCTCCAGTAAGTA	4750
AGCCGATTTTTGTTTACCTATGTAGGGACCTTGTACAAATATCTTGATAA	4800
CTGTGTTAGTCCAACTGGTAAGCGACTCTTAAGGAATTGGATCTGCCATC	. 4850
CACTCAAAGATGTAGAAAGCATCAATAAACGGCTTGATGTAGTTGAAGAA	4900
TTCACGGCAAACTCAGAAAGTATGCAAATCACTGGCCAGTATCTCCACAA	4950
ACTTCCAGACTTAGAAAGACTGCTCGGACGCATCAAGTCTAGCGTTCGAT	5000
CATCAGCCTCTGTTGCCTGCTCTTCTGGGGAAAAAAGTGCTGAAACAA	50 50
CGAGTAAGTATCACAAGTTTTCTGAGTAATGCCTTCCATGAGTAGT	5100
ATAGGACTAAAACATTACGGGTCTAGCTAAAGACTGTTCTCCTTCTTTTG	5150
CAATGTCTGGTTATTCATTACATTTCTCTTAACTTATTGCATTGCAGGTT	5200
AAAGCATTTGGGCAAATTGTGAAAGGGTTCAGAAGTGGAATTGATCTGTT	5250
GTTGGCTCTACAGAAGGAATCAAATATGATGAGTTTGCTTTATAAACTCT	5300
GTAAACTTCCTATATTAGTAGGAAAAAGCGGGCTAGAGTTATTTCTTTC	5350
CAATTCGAAGCAGCCATAGATAGCGACTTTCCAAATTATCAGGTGCCCAT	5400
CTATCTTTCATACTTTACAACAAAATGTCTGTCACTACTCAAAGCAATGC	5450
ATATGGCTTAGATCTCAACTCACCCCGAGGATCCTAAAGGGATTTGCT	5500
TTTTATTCCTAATGTTTTTGGATGGTTTGATTTATTTCTAACTTGAACTT	5550
ATTAATCTTGTACCAGAACCAAGATGTGACAGATGAAAACGCTGAAACTC	5600
TCACAATACTTATCGAACTTTTTATCGAAAGAGCAACTCAATGGTCTGAG	5650
GTCATTCACACCATAAGCTGCCTAGATGTCCTGAGATCTTTTGCAATCGC	5700
AGCAAGTCTCTCTGCTGGAAGCATGGCCAGGCCTGTTATTTTTCCCGAAT	5750
CAGAAGCTACAGATCAGAATCAGAAAACAAAAGGGCCAATACTTAAAATC	5800
CAAGGACTATGGCATCCATTTGCAGTTGCAGCCGATGGTCAATTGCCTGT	5850
TCCGAATGATATACTCCTTGGCGAGGCTAGAAGAAGCAGTGGCAGCATTC	5900
ATCCTCGGTCATTGTTACTGACGGGACCAAACATGGGCGGAAAATCAACT	5950
CTTCTTCGTGCAACATGTCTGGCCGTTATCTTTGCCCAAGTTTGTATACT	6000
CGTTAGATAATTACTCTATTCTTTGCAATCAGTTCTTCAACATGAATAAT	6050
AAATTCTGTTTTCTGTCTGCAGCTTGGCTGCTACGTGCCGTGTGAGTCTT	6100
GCGAAATCTCCCTCGTGGATACTATCTTCACAAGGCTTGGCGCATCTGAT	6150
AGAATCATGACAGGAGAGAGTAAGTTTTGTTCTCAAAATACCAATTCCTC	6200
GAACTATTTACTCAGATTTTGTCTGATTGGACAAGGTGGTTTTGCTTTTT	6250
TTTAGGTACCTTTTTGGTAGAATGCACTGAGACAGCGTCAGTTCTTCAGA	6300
ATGCAACTCAGGATTCACTAGTAATCCTTGACGAACTGGGCAGAGGAACT	6350
AGTACTTTCGATGGATACGCCATTGCATACTCGGTAACCTGCTCTTCTCC	6400
TTCAACTTATACTTGTTGATCAACAAAAACATGCAATTCATTTTGCTGAA	6450
ACTTATTGATTTATATCAGGTTTTTCGTCACCTGGTAGAGAAAGTTCAAT	6500
GTCGGATGCTCTTTGCAACACATTACCACCCTCTCACCAAGGAATTCGCG	6550
TCTCACCCACGTGTCACCTCGAAACACATGGCTTGCGCATTCAAATCAAG	6600

ATCTGATTATCAACCACGTGGTTGTGATCAAGACCTAGTGTTCTTGTACC	6650
GTTTAACCGAGGGAGCTTGTCCTGAGAGCTACGGACTTCAAGTGGCACTC	6700
ATGGCTGGAATACCAAACCAAGTGGTTGAAACAGCATCAGGTGCTGCTCA	6750
AGCCATGAAGAGATCAATTGGGGAAAACTTCAAGTCAAG	6800
CTGAGTTCTCAAGTCTGCATGAAGACTGGCTCAAGTCATTGGTGGGTATT	6850
TCTCGAGTCGCCCACAACAATGCCCCCATTGGCGAAGATGACTACGACAC	6900
TTTGTTTTGCTTATGGCATGAGATCAAATCCTCTTACTGTGTTCCCAAAT	6950
AAATGGCTATGACATAACACTATCTGAAGCTCGTTAAGTCTTTTGCTTCT	7000
CTGATGTTTATTCCTCTTAAAAAATGCTTATATATCAAAAAATTGTTTCC	7050
TCGATTATAACAAGATTATATGTATCTGTCGGTTTAGCTATGGTATAT	7100
AATATATGTATGTTCATGAGATTGGTCAAGAGAAATACTCACAAACAGTA	7150
TATTAAGAAGGAAATATGTTTATGCATTAATTTAAGTTTCAAGATAAACT	7200
GCAAATAACCTCGACTAAAGTTGCAAAGACCAAACACAAATTACAAAACT	7250
TATAAGACTTAAGTTCTGAATTCCCTAAAACCAAAAAAAA	7300
TATTTTGTTGCATCTACAAACAACACAAACCTACATAGTTTATAACTTAC	7350
TCATCACTGAGATTAACATCAGAATCATTCTCCATTTCTTCATCTTCACT	7400
CTCATCATCACCACCACCATGATGATTCTCCTCCTCTTCACGTAACC	7450
TAGCAATCTCACTCTGAGCTCTATCAACAATCTGCTTCTTCTGCAACTCC	7500
AAATCTCTCTGAAAATCAGCTCTCATCTTCTCCAACTCCTTCATTTGCTC	7550
TTTCTTACTCTTCTCCATCTTCTCATAAACCTTCCCAAACCTCTCAACAG	7,600
AATCCGCCAACATCTTATACGAAGCAGCGTCATTAACCTTCTTCCTCTCG	7650
TACTCAACCTCATCCTCATCCTCCTCCTCTCAGAATCACCAGGACT	<u>ु</u> 7700
ATCCATCATCATCAAACCCATTAGACTTATCTAAATAAACCTTAGTGT	7750
TCATAAACACAAACTCACCTGAATCAACACCACAAGCTAAACCTAAATCC	7800
GACTTGGGCGAAACACAAAGCAACATATCCAACTTATTGAAAAACGACCA	_7850
TTTACTTGAACCTAAACCTGATTTCTCAACCTTAATCTTCTCTTTTCTAT	7900
ACTTCCTCTTCAAGTCATCAATCATTCTCCTACATTGCGTCTCAGATTTC	7950
TCCATCCTTAGCTCCTCACTCACTTCTCAGCTACTTCATTCCAATCCTC	8000
GTTCCTCAAACTCCTTCTACCCAATTGCAAAAACCTATCTCCCCAAACTT	8050
CAAGCAACACAA	8062
•	

Figure 11 (Continued)



Comments/References: AtMSH3 3' side antisense : AtMSH3 3' (13 = 2104bp) from pUC18/13 Sal1/Sst1/T4 into pCW164 BamH1/T4 in Agrobacterium LBA44O4

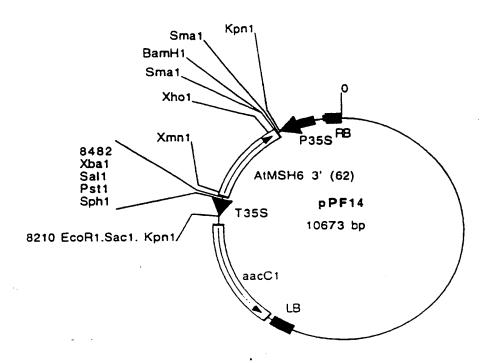


Figure 13

Comments/References: AtMSH6 (S8) 3' side antisens : 62 Sal1/Sst1/T4 (1379bp) into pCW164 BamH1/T4

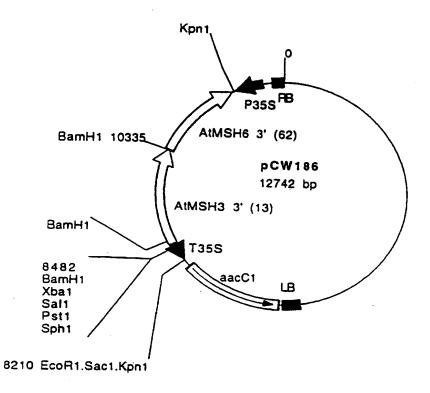


Figure 14

Comments/References: AtMSH6 3'/AtMSH3 3' antisense : AtMSH6 (S8) 3' side (62=1379bp) Sal1/Sst1/T4 into pPF13 (pCW164 AtMSH3 (S5) 3' side (13=2104) antisens)/Sma1. in LBA4404

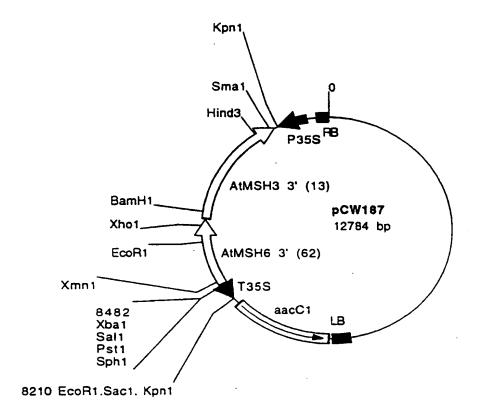


Figure 15

Comments/References: AtMSH3 3'/AtMSH6 3' antisens (D): AtMSH3 (S5) 3' side (13=2104bp) Sal1/Sst1/T4 into pPF14 (AtMSH6 (S8) 3'side (62=1379bp) antisense into pCW164)/Sma1. in LBA4404

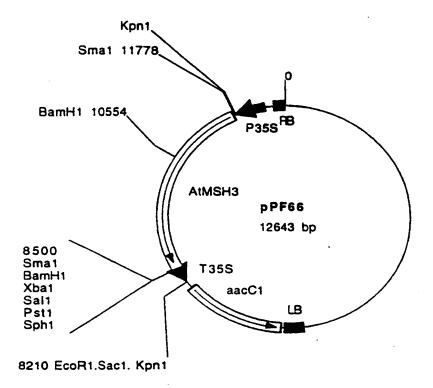
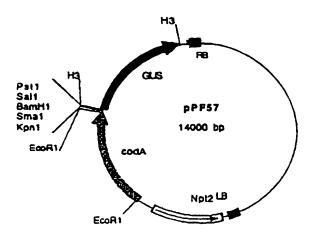
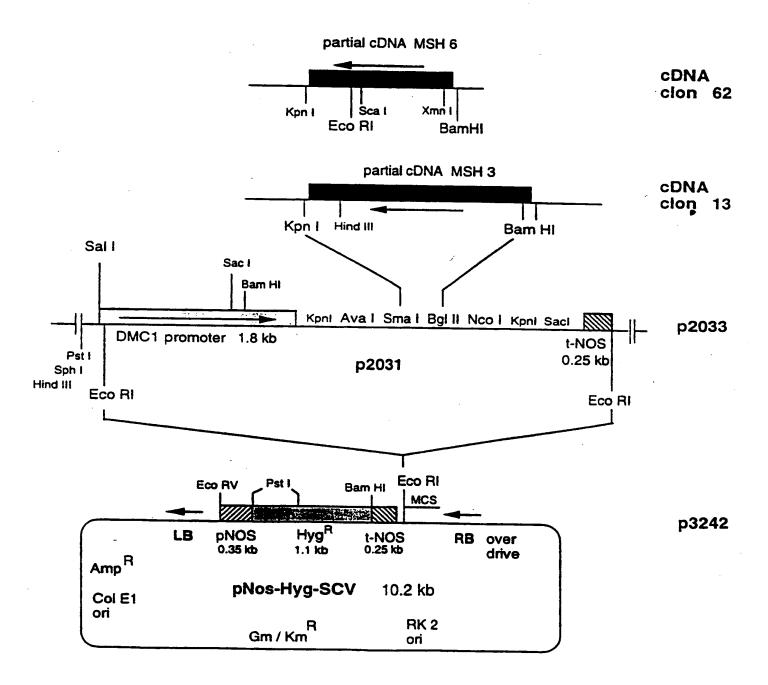


Figure 16

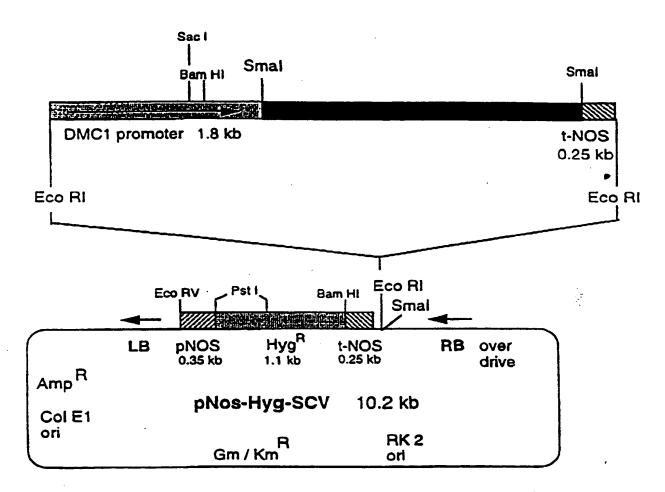
Comments/References: AtMSH3 (S8) complete, sense orientation : pPF26 (3342bp) Sma1 into pCW164 Sma1



Comments/References: pPZP111 with codA EcoR1 cassette in EcoR1 site and HInd3 GUS cassette in Hind3 site. KanR. All genes under Promoter/terminator 35S



p3243



فالموارد فالمعارب والمحافظ

1

SEQUENCE LISTING

```
<110>
                Rhone-Poulenc Agro; Betzner, Andreas Stefan; Doutriaux,
               Marie-Pascale; Freyssinet, Georges; Perez, Pascual.
               Methods for obtaining plant varieties
  <120>
  <130>
               395498C
  <150>
               PO9745
  <151>
               1997-10-10
  <160>
               98
  <210>
               1
  <211>
               23
  <212>
               DNA
 <213>
              Artificial sequence
 <220>
 <221>
              modified_base
 <222>
              11
 <223>
              I
 <220>
 <221>
              modified_base
 <222>
              14
 <223>
              Ι
 <220>
 <221>
              modified_base
 <222>
              17
 <223>
              I
<220>
<223>
              Degenerate oligonucleotides UPMU used to isolate AtMSH3 and
<300>
<301>
             Reenan and Kolodner
<302>
             Genetics
<303>
             132
<306>
             963-973
<307>
             1992
<400>
             1
ctggatccac nggnccnaay atg
                                                                        23
```

<210>

<211>

<212>

2

23

DNA

WO 99/19492 PCT/EP98/06977

2

```
Artificial sequence
 <213>
 <220>
              modified base
 <221>
 <222>
              15
 <223>
              Ι
 <220>
              modified base
 <221>
 <222>
              18
 <223>
              I
 <220>
              Degenerate oligonucleotides DOMU used to isolate AtMSH3 and
 <223>
              AtMSH6.
 <300>
 <301>
              Reenan and Kolodner
 <302>
              Genetics
 <303>
              132
<306>
              963-973
 <307>
              1992
 <400>
 ctggatccrt artgngtnrc raa
                                                                        23
 <210>
              3
 <211>
              24
 <212>
              DNA
 <213>
              Artificial sequence
 <220>
              MSH3 specific primer 636 for PCR using cDNA of Arabidopsis
              thaliana ecotype Columbia
 <400>
                                                                        24
 tgctagtgcc tcttgcaagc tcat
 <210>
 <211>
              27
              DNA
 <212>
 <213>
              Artificial sequence
 <220>
              Primer AP1 for PCR using cDNA of Arabidopsis thaliana ecotype
 <223>
              Columbia containing adapter sequences ligated to both its
              ends
 <400>
```

ccatcctaat	acgactcact atagggc	27
<210>	5	
<211>	23	
<212>	DNA	
<213>	Artificial sequence	
<220>		
<223>	Primer AP2 for PCR using cDNA of Arabidopsis thaliana e Columbia containing adapter sequences ligated to both i ends	cotype
<400>	5	
actcactata	gggctcgagc ggc	23
<210>	6	
<211>	30	
<212>	DNA	•
<213>	Artificial sequence	
<220>		
<223>	MSH3 specific primer S525 for PCR using cDNA of Arabido thaliana ecotype Columbia	psis
<400>	6	
aggttctgat :	tatgtgtgac gctttactta	30
<210>		
<211>	7	
<212>	29	
<213>	DNA	
(213)	Artificial sequence	
<220>		
<223>	MSH3 specific primer S51 for PCR using cDNA of Arabidops thaliana ecotype Columbia	sis
<400>	7	
ggatcgggta c	tgggttttg agtgtgagg	29
<210>	8	
<211>	24	
<212>	DNA	
<213>	Artificial sequence	
•		•
<220>		
<223>	MSH3 specific primer 635 for PCR using cDNA of Arabidops thaliana ecotype Columbia	is

WO 99/19492 PCT/EP98/06977

< 400 >	8 ,	
gcacgtgctt	gatggtgttt tcac 24	Ŀ
• • •		
<210>	9	
<211>	28	
<212>	DNA	
<213> -	Artificial sequence	
<220>		
<223>	MSH3 specific primer S523 for PCR using cDNA of Arabidopsi thaliana ecotype Columbia	İS
<400>	9	
tcagacagta	tccagcatgg cagaagta 28	3
<210>	10	•
<211>	33	
<212>	DNA	
<213>	Artificial sequence	
<220>		
<223>	MSH3 specific primer 1S5 for PCR using cDNA of Arabidopsis thaliana ecotype Columbia	3
<400>	10	
atcccgggat	gggcaagcaa aagcagcaga cga 33	3
<210>	11	
<211>	27 DNA	
<212>		
<213>	Artificial sequence	
<220>		
<223>	MSH3 specific primer S53 for PCR using cDNA of Arabidopsisthaliana ecotype Columbia	3
<400>	11	
gacaaagagc	gaaatgaggc cccttgg 2	7
-210-		
<210>	12	
<211>	1250 DNA	
<212> <213>	DNA	
<213> <223>	Arabidopsis thaliana ecotype Columbia Clone 52	

<	4	0	0	>	1	2
---	---	---	---	---	---	---

cccgggatgg	gcaagcaaaa	gcagcagacg	atttctcgtt	tcttcgctcc	caaacccaaa	60
tccccgactc	acgaaccgaa	tccggtagcc	gaatcatcaa	caccgccacc	gaagatatcc	120
gccactgtat	ccttctctcc	ttccaagcgt	aagcttctct	ccgaccacct	cgccgccgcg	180
tcacccaaaa	agcctaaact	ttctcctcac	actcaaaacc	cagtacccga	tcccaattta	240
caccaaagat	ttctccagag	atttctggaa	ccctcgccgg	aggaatatgt	tcccgaaacg	300
tcatcatcga	ggaaatacac	accattggaa	cagcaagtgg	tggagctaaa	gagcaagtac	360
ccagatgtgg	ttttgatggt	ggaagttggt	tacaggtaca	gattcttcgg	agaagacgcg	420
gagatcgcag	cacgcgtgtt	gggtatttac	gctcatatgg	atcacaattt	catgacggcg	480
agtgtgccaa	catttcgatt	gaatttccat	gtgagaagac	tggtgaatgc	aggatacaag	540
attggtgtag	tgaagcagac	tgaaactgca	gccattaagt	cccatggtgc	aaaccggacc	6 00
ggcccttttt	tetggggaet	grcggcgrtg	tataccaaag	ccacgcttga	agcggctgag	660
gatataagtg	grggrrgrgg	tggtgaagaa	ggttttggtt	cacagagtaa	tttcttggtt	720
tgtgttgtgg	atgagagagt	taagtcggag	acattaggct	gtggtattga	aatgagtttt	780
gatgttagag	tcggtgttgt	tggcgttgaa	atttcgacag	gtgaagttgt	ttatgaagag	840
ttcaatgata	atttcatgag	aagtggatta	gaggctgtga	ttttgagctt	gtcaccagct	900
gagetgttge	ttggccagcc	tctttcacaa	caaactgaga	agtttttggt	ggcacatgct	960
ggacctacct	caaacgttcg	agtggaacgt	gcctcactgg	attgtttcag	caatggtaat	1020
gcagtagatg	aggttatttc	attatgtgaa	aaaatcagcg	caggtaactt	agaagatgat	1080
aaagaaatga	agctggaggc	tgctgaaaaa	ggaatgtctt	gcttgacagt	tcatacaatt	1140
atgaacatgc	cacatctgac	tgttcaagcc	ctcgccctaa	cgttttgcca	tctcaaacag	1200
tttggatttg	aaaggatcct	ttaccaaggg	gcctcatttc	gctctttgtc		1,250

<210> 13
<211> 34
<212> DNA
<213> Artificial sequence

<220>

<223> MSH3 specific primer 2S5 for PCR using cDNA of Arabidopsis thaliana ecotype Columbia

<4005	13					
accocgggcc	aaaatgaaca agttgg	sttt agtc			34	
<210>	14					
<211>	27					
<212>	DNA	107.50				
<213>	Artificial sequ	ience				
<220>						
<223>	MSW3 specific r	rimer S52 for PC	R using cDN	A of Arabido	nsi	s
~223/	thaliana ecotyp		ub 2 g		,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,	_
<400>	14					
gccacatctg	actgttcaag ccctc	gc .			27	
<210>						
<211>	2110					₽
<212>	DNA				•	•
<213>	Arabidopsis the	aliana ecotype Co	olumbia			
<223>	Clone 13				ŕ	
<400>	15				. Arr	
gccacatctg	actgttcaag ccctc	gccct aacgttttgc	catctcaaac	agtttggatt	7.	60
tgaaaggatc	ctttaccaag gggcc	tcatt tcgctctttg	tcaagtaaca	cagagatgac		120
teteteagee	aatactctgc aacag	ttgga ggttgtgaaa	aataattcag	atggatcgga		180
atctggctcc	ttattccata atatg	aatca cacacttaca	gtatatggtt	ccaggcttct		240
tagacactgg	gtgactcatc ctcta	tgcga tagaaatttg	atatctgctc	ggcttgatgc		300
tgtttctgag	atttctgctt gcatg	ggatc tcatagttct	tcccagetca	gcagtgagtt		360
ggttgaagaa	ggttctgaga gagca	attgt atcacctgag	ttttatctcg	tgctctcctc		420
agtcttgaca	gctatgtcta gatca	tctga tattcaacgt	ggaataacaa	gaatctttca		480
tcggactgct	aaagccacag agtto	attgc agttatggaa	gctattttac	ttgcggggaa	٠	540
gcaaattcag	cggcttggca taaag	caaga ctctgaaatg	aggagtatgc	aatctgcaac		600
tgtgcgatct	actettttga gaaaa	ittgat ttctgttatt	tcatcccctg	ttgtggttga		660
caatgccgga	aaacttctct ctgcc	ctaaa taaggaagcg	gctgttcgag	gtgacttgct	*	720
cgacatacta	atcacttcca gcgad	caatt teetgagett	gctgaagctc	gccaagcagt		780
tttagtcatc	agggaaaagc tggat	tcctc gatagettca	tttcgcaaga	agctcgctat		840

togaaatttg	gaatttcttc	aagtgtcggg	gatcacacat	ttgatagagc	tgcccgttga	900
ttccaaggtc	cctatgaatt	gggtgaaagt	aaatagcacc	aagaagacta	ttcgatatca	960
tcccccagaa	atagtagctg	gcttggatga	gctagctcta	gcaactgaac	atcttgccat	1020
tgtgaaccga	gcttcgtggg	atagtttcct	caagagtttc	agtagatact	acacagattt	1080
taaggctgcc	gttcaagctc	ttgctgcact	ggactgtttg	cactcccttt	caactctatc	1140
tagaaacaag	aactatgtcc	gtcccgagtt	tgtggatgac	tgtgaaccag	ttgagataaa	1200
catacagtct	ggtcgtcatc	ctgtactgga	gactatatta	caagataact	tcgtcccaaa	1260
tgacacaatt	ttgcatgcag	aaggggaata	ttgccaaatt	atcaccggac	ctaacatggg	1320
aggaaagagc	tgctatatcc	gtcaagttgc	tttaatttcc	ataatggctc	aggttggttc	1380
ctttgtacca	gcgtcattcg	ccaagctgca	cgtgcttgat	ggtgttttca	ctcggatggg	1440
tgcttcagac	agtatccagc	atggcagaag	tacctttcta	gaagaattaa	gtgaagcgtc	1500
acacataatc	agaacctgtt	cttctcgttc	gcttgttata	ttagatgagc	ttggaagagg	1560
cactagcaca	cacgacggtg	tagccattgc	ctatgcaaca	ttacagcatc	tcctagcaga	1620
aaagagatgt	ttggttcttt	ttgtcacgca	ttaccctgaa	atagctgaga	tcagtaacgg	1680
attcccaggt	tctgttggga	cataccatgt	ctcgtatctg	acattgcaga	aggataaagg	1740
cagttatgat	catgatgatg	tgacctacct	atataagctt	gtgcgtggtc	tttgcagcag	1800
gagctttggt	tttaaggttg	ctcagcttgc	ccagatacct	ccatcatgta	tacgtcgagc	1860
catttcaatg	gctgcaaaat	tggaagctga	ggtacgtgca	agagagaa	atacacgcat	1920
gggagaacca	gaaggacatg	aagaaccgag	aggcgcagaa	gaatctattt	cggctctagg	1980
tgacttgttt	gcagacctga	aatttgctct	ctctgaagag	gacccttgga	aagcattcga	2040
gtttttaaag	catgcttgga	agattgctgg	caaaatcaga	ctaaaaccaa	cttgttcatt	2100
ttgacccggg						2110

<210>	16
<211>	29
<212>	DNA
<213>	Artificial sequence
<220>	
<223>	MSH3 specific primer S51 for PCR using cDNA of Arabidopsis thaliana ecotype Columbia
<400>	16

ggat	cggg	rta c	tggg	rttt	g ag	tgtg	agg									2 9
<210			17													
<211			30													
<212			DNA													
<213	>		Art	ific	ial	sequ	ence	:								
<220	>															
<223	>				ecif a ec					or P	CR u	sing	CDN	IA of	Arab	oidopsis
< 400	>		17													
aggt	tctg	at t	atgt	gtga	ic go	ttta	CCCE	1								3 (
<210	>		18													
<211	>		352	2												
<212	>		DNA													
<213	>		Ara	bido	psis	: cha	iia:	a ec	otyp	e Co	lumi	ia				
<220	· >										•					?
<221	.>		CDS	;												
<222					(3	342)										**
<223								CDNA	and	l ded	luces	. sec	ruènc	e of	the	encoded
		•			tide		-						-			·•
< 400	>		18						•							rs.
ccta	agaa	ag d	gege	gaaa	a tt	ggca	acco	c aag	jtte	gcca	tago	cac	gac o	cacga	acctt	. 60
catt	tctc	ett a	aacq	gagg	ga ga	ttac	gaat	aaa	agcaa	act						99
٠																
_		_		_	_	-	_			_			_	ccc		147
	Gly	Lys	Gln		Gln	Gln	Thr	Ile		Arg	Phe	Phe	Ala	Pro	Lys	
. 1				5					10					15		
ccc	aaa	tcc	ccg	act	cac	gaa	ccg	aat	ccg	gta	gcc	gaa	tca	tca	aca	195
Pro	Lys	Ser	Pro	Thr	His	Glu	Pro	Asn	Pro	Val-	Ala	Glu	Ser	Ser	Thr	
			20					25					30			
cca	сса	cca	aacr	ata	tcc	acc	act	qta	tcc	ttc	tct	cct	tcc	aag	cat	243
_		_	_			_		_						Lys	-	
		35	-1-				40					45		-1-	5	
226		a= a		~~~		a.		~~~	~~~							291
														CCL		291
nys	50	red	ser	Asp	uis	55	ALG	AIG	AIG	ser	60	rys	гÀг	Pro	rys	
~++	F.C.F	cct	C2.C	a.c.+	C22	220	CCS	~+ =	666	<i>(</i> 73 ⊢	CC-	225	++ =	cac	Cae	339
														His		,,,,
65	3 C L	-10	HIS	TILL	70	VOII	FIO	AGT	FIO	_	FIO	WOII	neu	ura		
0 0					, 0					75					80	

																,	
aga Arg	ttt Phe	ctc Leu	cag Gln	aga Arg 85	Pne	ctg Leu	gaa Glu	ccc Pro	ser 90	Pro	gag Glu	gaa Glu	tat Tyr	gtt Val 95	ccc Pro	38	7
gaa Glu	acg Thr	cca Ser	ser	ser	agg Arg	aaa Lys	tac Tyr	aca Thr 105	Pro	ttg Leu	gaa Glu	cag Gln	caa Gln 110	Val	gtg Val	43	5
gag Glu	cta Leu	aag Lys 115	ser	aag Lys	tac Tyr	cca Pro	gat Asp 120	gtg Val	gtt Val	ttg Leu	atg Met	gtg Val 125	gaa Glu	gtt Val	ggt Gly	48	3
tac Tyr	agg Arg 130	Tyr	aga Arg	ttc Phe	ttc Phe	gga Gly 135	gaa Glu	gac Asp	gcg Ala	gag Glu	atc Ile 140	gca Ala	gca Ala	cgc Arg	gtg Val	53	1
ttg Leu 145	ggt	att Ile	tac Tyr	gct Ala	cat His 150	atg Met	gat Asp	cac His	aat Asn	ttc Phe 155	atg Met	acg Thr	gcg Ala	agt Ser	gtg Val 160	57	9
cca Pro	aca Thr	ttt Phe	cga Arg	ttg Leu 165	aat Asn	ttc Phe	cat His	gtg Val	aga Arg 170	aga Arg	ctg Leu	gtg Val	aat Asn	gca Ala 175	gga Gly	62	7
tac Tyr	aag Lys	att Ile	ggt Gly 180	gta Val	gtg Val	aag Lys	cag Gln	act Thr 185	gaa Glu	act Thr	gca Ala	gcc Ala	att Ile 190	aag Lys	tcc Ser	675	5
cat His	ggt Gly	gca Ala 195	aac Asn	cgg Arg	acc Thr	ggc Gly	cct Pro 200	ttt Phe	ttc Phe	cgg Arg	gga Gly	ctg Leu 205	tcg Ser	gcg Ala	ttg Leu	723	3
tat Tyr	acc Thr 210	aaa Lys	gcc Ala	acg Thr	ctt Leu	gaa Glu 215	gcg Ala	gct Ala	gag Glu	gat Asp	ata Ile 220	agt Ser	ggt Gly	ggt Gly	tgt Cys	77;	L
ggt Gly 225	ggt Gly	gaa Glu	gaa Glu	ggt Gly	ttt Phe 230	ggt Gly	tca Ser	cag Gln	agt Ser	aat Asn 235	ttc Phe	ttg Leu	gtt Val	tgt Cys	gtt Val 240	819	,
gtg Val	gat Asp	gag Glu	aga Arg	gtt Val 245	aag Lys	tcg Ser	gag Glu	aca Thr	tta Leu 250	ggc Gly	tgt Cys	ggt Gly	att Ile	gaa Glu 255	atg Met	867	,
agt Ser	ttt Phe	gat Asp	gtt Val 260	aga Arg	gtc Val	ggt Gly	gtt Val	gtt Val 265	ggc Gly	gtt Val	gaa Glu	att Ile	tcg Ser 270	aca Thr	ggt Gly	915	; •
gaa Glu	gtt Val	gtt Val 275	tat Tyr	gaa Glu	gag Glu	ttc Phe	aat Asn 280	gat Asp	aat Asn	ttc Phe	atg Met	aga Arg 285	agt Ser	gga Gly	tta Leu	963	

														ggc Gly		1011
														gga Gly		1059
			_	_			_	_		_	_	-		agc Ser 335		1107
		-	-	_		_					_			agc Ser	-	1155
			_	_	_					-		_		gaa Glu		1203
			_	_										cat His	_	1251
	_		-		_		_		_				_	ttt Phe		1299
	_							_			_		_	cca Ser 415	_	1347
			_				-			-		_	_	gag Glu	•	1395
-					_	_		_						cat His		1443
-						-		_			•		_	cac His	tgg Trp	1491
											Ser			ctt Leu		1539
-	_						_	_		Ser		_		tcc Ser 495	Gln	1587

11	
ctc agc agt gag ttg gtt gaa gaa ggt tct gag aga gca att gta tca Leu Ser Ser Glu Leu Val Glu Glu Gly Ser Glu Arg Ala Ile Val Ser 500 505 510	1635
cct gag ttt tat ctc gtg ctc tcc tca gtc ttg aca gct atg tct aga Pro Glu Phe Tyr Leu Val Leu Ser Ser Val Leu Thr Ala Met Ser Arg 515 520 525	1683
Ser Ser Asp Ile Gln Arg Gly Ile Thr Arg Ile Phe His Arg Thr Ala 530 535 540	1731
aaa gcc aca gag ttc att gca gtt atg gaa gct att tta ctt gcg ggg Lys Ala Thr Glu Phe Ile Ala Val Met Glu Ala Ile Leu Leu Ala Gly 545 550 560	1779
aag caa att cag cgg ctt ggc ata aag caa gac tct gaa atg agg agt Lys Gln Ile Gln Arg Leu Gly Ile Lys Gln Asp Ser Glu Met Arg Ser 565 570 575	1827
atg caa tot goa act gtg oga tot act ott ttg aga aaa ttg att tot Met Gln Ser Ala Thr Val Arg Ser Thr Leu Leu Arg Lys Leu Ile Ser 580 585 590	1875
gtt att toa toe oot gtt gtg gtt gao aat goo gga aaa ott oto tot Val Ile Ser Ser Pro Val Val Val Asp Asn Ala Gly Lys Leu Leu Ser 595 600 605	1923
gcc cta aat aag gaa gcg gct gtt cga ggt gac ttg ctc gac ata cta Ala Leu Asn Lys Glu Ala Ala Val Arg Gly Asp Leu Leu Asp Ile Leu 610 615 620	1971
atc act tcc agc gac caa ttt cct gag ctt gct gaa gct cgc caa gca Ile Thr Ser Ser Asp Gln Phe Pro Glu Leu Ala Glu Ala Arg Gln Ala 625 630 635 640	2019
gtt tta gtc atc agg gaa aag ctg gat tcc tcg ata gct tca ttt cgc Val Leu Val Ile Arg Glu Lys Leu Asp Ser Ser Ile Ala Ser Phe Arg 645 650 655	2067
aag aag ctc gct att cga aat ttg gaa ttt ctt caa gtg tcg ggg atc Lys Lys Leu Ala Ile Arg Asn Leu Glu Phe Leu Gln Val Ser Gly Ile 660 665 670	2115
aca cat ttg ata gag ctg ccc gtt gat tcc aag gtc cct atg aat tgg Thr His Leu Ile Glu Leu Pro Val Asp Ser Lys Val Pro His Asn Trp 675 680 685	2163
gtg aaa gta aat agc acc aag aag act att cga tat cat ccc cca gaa Val Lys Val Asn Ser Thr Lys Lys Thr Ile Arg Tyr His Pro Pro Glu 690 695 700	2211

ata Ile 705	gta Val	gct Ala	ggc Gly	ttg Leu	gat Asp 710	gag Glu	cta Leu	gct Ala	cta Leu	gca Ala 715	act Thr	gaa Glu	cat His	ctt Leu	gcc Ala 720	2259
att Ile	gtg Val	aac Asn	cga Arg	gct Ala 725	tcg Ser	tgg Trp	gat Asp	agt Ser	ttc Phe 730	ctc Leu	aag Lys	agt Ser	ttc Phe	agt Ser 735	aga Arg	2307
														ctg Leu		2355
tgt Cys	ttg Leu	cac His 755	tcc Ser	ctt Leu	tca Ser	act Thr	cta Leu 760	tct Ser	aga Arg	aac Asn	aag Lys	aac Asn 765	tat Tyr	gtc Val	cgt Arg	2403
														cag Gln		2451
														gtc Val		249\$
														atc Ile 815		2547
														gct Ala		2595
														ttc Phe		2643
												Gly	-	tca Ser	_	2691
														gaa Glu		2739
														tta Leu 895		2787
														gcc Ala		2835

gca Ala	aca Thr	tta Leu 915	cag Gln	cat His	ctc Leu	cta Leu	gca Ala 920	gaa Glu	aag Lys	aga Arg	tgt Cys	ttg Leu 925	gtt Val	ctt Leu	ttt Phe	2883
gtc Val	acg Thr 930	cat His	tac Tyr	cct Pro	gaa Glu	ata Ile 935	gct Ala	gag Glu	atc Ile	agt Ser	aac Asn 940	gga Gly	ttc Phe	cca Pro	ggt Gly	2931
tct Ser 945	gtt Val	GJÅ aaa	aca Thr	tac Tyr	cat His 950	gtc Val	tcg Ser	tat Tyr	ctg Leu	aca Thr 955	ttg Leu	cag Gln	aag Lys	gat Asp	aaa Lys 960	2979
Gly	agt Ser	tat Tyr	gat Asp	cat His 965	gat Asp	gat Asp	gtg Val	acc Thr	tac Tyr 970	cta Leu	tat Tyr	aag Lys	ctt Leu	gtg Val 975	cgt Arg	3027
ggt Gly	ctt Leu	tgc Cys	agc Ser 980	agg Arg	agc Ser	ttt Phe	ggt Gly	ttt Phe 985	aag Lys	gtt Val	gct Ala	cag Gln	ctt Leu 990	gcc Ala	cag Gln	3075
ata Ile	cct Pro	cca Pro 995	tca Ser	tgt Cys	ata Ile	Arg	cga Arg .000	gcc Ala	att Ile	tca Ser	Met	gct Ala .005	gca Ala	aaa Lys	ttg Leu	3123
GIU	gct Ala .010	gag Glu	gta Val	cgt Arg	Ala	aga Arg .015	gag Glu	aga Arg	aat Asn	Thr	cgc Arg .020	atg Met	gga Gly	gaa Glu	cca Pro	3171
gaa Glu 1025	GIY	cat	gaa Glu	Glu	ccg Pro .030	aga Arg	ggc	gca Ala	Glu	gaa Glu 035	tct Ser	att Ile	tcg Ser	Ala	cta Leu 040	3219
ggt Gly	gac Asp	ttg Leu	Phe	gca Ala 045	gac Asp	ctg Leu	aaa Lys	Phe	gct Ala 050	ctc Leu	tct Ser	gaa Glu	Glu	gac Asp .055	cct Pro	3267
tgg Trp	aaa Lys	gca Ala 1	ttc Phe 060	gagʻ Glu	ttt Phe	tta Leu	Lys	cat His 065	gct Ala	tgg Trp	aag Lys	Ile	gct Ala 070	ggc Gly	aaa Lys	3315
atc Ile	Arg	cta Leu 075	aaa Lys	cca Pro	act Thr	Cys	tca Ser 080	ttt Phe	tgat	ttaa	țc t	taac	atta	t		3362
agca	actg	ca a	ggtc	ttga	t ca	tctg	ttag	ttg	cgta	cta	actt	atgt	gt a	ttag	tataa	3422
caag	aaaa	ga g	aatt	agag	a ga	tgga	ttct	aat	ccgg	tgt	tgca	gtac	at c	ttt	ctcca	3482
cccg	cata	aa a	aaaa	aaaa	a aa	aaaa	aaaa	aaa	aaaa	aaa						3522
<210	>		19				•									

<212> PRT

1081

<211>

<213> Arabidopsis thaliana ecotype Columbia <223> Polypeptide MSH3 <400> 19

Met Gly Lys Gln Lys Gln Gln Thr Ile Ser Arg Phe Phe Ala Pro Lys
1 5 10 15

Pro Lys Ser Pro Thr His Glu Pro Asn Pro Val Ala Glu Ser Ser Thr
20 25 30

Pro Pro Pro Lys Ile Ser Ala Thr Val Ser Phe Ser Pro Ser Lys Arg
35 40 45

Lys Leu Leu Ser Asp His Leu Ala Ala Ser Pro Lys Lys Pro Lys 50 55 60

Leu Ser Pro His Thr Gln Asn Pro Val Pro Asp Pro Asn Leu His Gln 65 70 75 80

Arg Phe Leu Gln Arg Phe Leu Glu Pro Ser Pro Glu Glu Tyr Val Pro
85 90 95

Glu Thr Ser Ser Ser Arg Lys Tyr Thr Pro Leu Glu Gln Gln Val Val
100 105 110

Glu Leu Lys Ser Lys Tyr Pro Asp Val Val Leu Met Val Glu Val Gly
115 120 125

Tyr Arg Tyr Arg Phe Phe Gly Glu Asp Ala Glu Ile Ala Ala Arg Val 130 135 140

Leu Gly Ile Tyr Ala His Met Asp His Asn Phe Met Thr Ala Ser Val 145 150 155 160.

Pro Thr Phe Arg Leu Asn Phe His Val Arg Arg Leu Val Asn Ala Gly
165 170 175

Tyr Lys Ile Gly Val Val Lys Gln Thr Glu Thr Ala Ala Ile Lys Ser 180 185 190

His Gly Ala Asn Arg Thr Gly Pro Phe Phe Arg Gly Leu Ser Ala Leu 195 200 205

Tyr Thr Lys Ala Thr Leu Glu Ala Ala Glu Asp Ile Ser Gly Gly Cys
210 215 220

Gly Gly Glu Glu Gly Phe Gly Ser Gln Ser Asn Phe Leu Val Cys Val 225 230 235 240

Val Asp Glu Arg Val Lys Ser Glu Thr Leu Gly Cys Gly Ile Glu Met 245 6 250 255

Ser	Phe	Asp	Val 260	Arg	Val	Gly	Val	Val 265	Gly	Val	Glu	Ile	Ser 270	Thr	Gly
Glu	Val	Val 275	Tyr	Glu	Glu	Phe	Asn 280	Asp	Asn	Phe	Met	Arg 285	Ser	Gly	Leu
Glu _,	Ala 290	Val	Ile	Leu	Ser	Leu 295	Ser	Pro	Ala	Glu	Leu 300	Leu	Leu	Gly	Gln
Pro 305	Leu	Ser	Gln	Gln	Thr 310	Glu	Lys	Phe	Leu	Val 315	Ala	Met	Ala	Gly	Pro 320
Thr	Ser	Asn	Val	Arg 325	Val	Glu	Arg	Ala	Ser 330	Leu	Asp	Cys	Phe	Ser 335	Asn
Gly	Asn	Ala	Val 340	Asp	Glu	Val	Ile	Ser 345	Leu	Cys	Glu	Lys	Ile 350	Ser	Ala
Gly	Asn	Leu 355	Glu	Asp	Asp	Lys	Glu 360	Met	Lys	Leu	Glu	Ala 365	Ala	Glu	Lys
Gly	Met 370	Ser	CÀR	Leu	Thr	Val 375	His	Thr	Ile	Met	Asn 380	Met	Pro	His	Leu
Thr 385	Val	Gln	Ala	Leu	Ala 390	Leu	Thr	Phe	Cys	His 395	Leu	Lys	Gln	Phe	Gly 400
Phe	Glu	Arg	Ile	Leu 405	Tyr	Gln	Gly	Ala	Ser 410	Phe	Arg	Ser	Leu	Ser 415	Ser
Asn	Thr	Glu	Met 420	Thr	Leu	Ser	Ala	Asn 425	Thr	Leu	Gln	Gln	Leu 430	Glu	Val
Val	Lys	Asn 435	Asn	Ser	Asp	Gly	Ser 440	Glu	Ser	Gly	Ser	Leu 445	Phe	His	Asn
Met	Asn 450	His	Thr	Leu	Thr	Val 455	Tyr	Gly	Ser	Arg	Leu 460	Leu	Arg	His	Trp
Val 465	Thr	His	Pro	Leu	Cys 470	Asp	Arg	Asn	Leu	Ile 475	Ser	Ala	Arg	Leu	Asp 480
Ala	Val	Ser	Glu	Ile 485	Ser	Ala	Cys	Met	Gly 490	Ser	His	Ser	Ser	Ser 495	Glr
Leu	Ser	Ser	Glu 500	Leu	Val	Glu	Glu	Gly 505	Ser	Glu	Arg	Ala	Ile 510	Val	Ser
Pro	Glu	Phe 515	туr	Leu	Val	Leu	Ser 520	Ser	Val	Leu	Thr	Ala 525	Met	Ser	Arg
Ser	Ser 530	Asp	Ile	Gln	Arg	Gly 535	Ile	Thr	Arg	Ile	Phe 540	His	Arg	Thr	Ala

WO 99/19492 PCT/EP98/06977

Lys 545		Thr	Glu	Phe	Ile 550	Ala	Val	Met	Glu	Ala 555	Ile	Leu	Leu	Ala	Gly 560
Lys	Gln	Ile	Gln	Arg 565	Leu	Gly	Ile	Lys	Gln 570	Asp	Ser	Glu	Met	Arg 575	Ser
Met	Gln	Ser	Ala 580	Thr	Val	Arg	Ser	Thr 585	Leu	Leu	Arg	Lys	Leu 590	Ile	Ser
Val	Ile	Ser 595	Ser	Pro	Val	Val	Val 600	Asp	Asn	Ala	Gly	Lys 605	Ļeu	Leu	Ser
Ala	Leu 610	Asn	Lys	Glu	Ala	Ala 615	Val	Arg	Gly	Asp	Leu 620	Leu	Asp	Ile	Leu
Ile 625	Thr	Ser	Ser	Asp	Gln 630	Phe	Pro	Glu	Leu	Ala 635	Glu	Ala	Arg	Gln	Ala 640
Val	Leu	Val	Ile	Arg 645	Glu	Lys	Leu	Asp	Ser 650	Ser	Ile	Ala	Ser	Phe 655	Arg
Lys	Lvs	Leu	Ala 660	Ile	Arg	Asn	Leu	Glu 665	Phe	Leu	Gln	Val	Ser 670	Gly	Ile
Thr	His	Leu 675	Ile	Glu	Leu	Pro	Val 680	Asp	Ser	Lys	Val	Pro 685	His	Asn	Trp
Val	Lys 690	Val	Asn	Ser	Thr	Lys 695	Lys	Thr	Ile	Arg	Tyr 700	His	Pro	Pro	Glu
Ile 705	Val	Ala	Gly	Leu	Asp 710	Glu	Leu	Ala	Leu	Ala 715	Thr	Glu	His	Leu	Ala 720
Ile	Val	Asn	Arg	Ala 725	Ser	Trp	Asp	Ser	Phe 730	Leu	Lys	Ser	Phe	Ser 735	Arg
Tyr	Tyr	Thr	Asp 740	Phe-	Lys	Ala	Ala	Val 745		Ala	Leu	Ala	Ala 750	Leu	Asp
Cys	Leu	His 755	Ser	Leu	Ser	Thr	Leu 760	Ser	Arg	Asn	Lys	Asn 765	Tyr	Val	Arg
Pro	Glu 770	Phe	Val	Asp	Asp	Cys 775	Glu	Pro	Val	Glu	Ile 780	Asn	Ile	Gln	Ser
Gly 785	Arg	His	Pro	Val	Leu 790	Glu	Thr	Ile	Leu	Gln 795	Asp	Asn	Phe	Val	Pro 800
Asn	Asp	Thr	Ile	Leu 805	His	Ala	Glu	Gly	Glu 810	Tyr	Суѕ	Gln	Ile	Ile 815	Thr
Gly	Pro	Asn	Met 820	Gly	Gly	Lys	Sar	Cys 825	Tyr	Ile	Arg	Gln	Val 830	Ala	Leu

Ile Ser Ile Met Ala Gln Val Gly Ser Phe Val Pro Ala Ser Phe Ala 835 840 845

Lys Leu His Val Leu Asp Gly Val Phe Thr Arg Met Gly Ala Ser Asp 850 855 860

Ser Ile Gln His Gly Arg Ser Thr Phe Leu Glu Glu Leu Ser Glu Ala 865 870 875 880

Ser His Ile Ile Arg Thr Cys Ser Ser Arg Ser Leu Val Ile Leu Asp 885 890 895

Glu Leu Gly Arg Gly Thr Ser Thr His Asp Gly Val Ala Ile Ala Tyr 900 905 910

Ala Thr Leu Gln His Leu Leu Ala Glu Lys Arg Cys Leu Val Leu Phe 915 920 925

Val Thr His Tyr Pro Glu Ile Ala Glu Ile Ser Asn Gly Phe Pro Gly 930 935 940

Ser Val Gly Thr Tyr His Val Ser Tyr Leu Thr Leu Gln Lys Asp Lys 945 950 955 960

Gly Ser Tyr Asp His Asp Asp Val Thr Tyr Leu Tyr Lys Leu Val Arg 965 970 975

Gly Leu Cys Ser Arg Ser Phe Gly Phe Lys Val Ala Gln Leu Ala Gln 980 985 990

Ile Pro Pro Ser Cys Ile Arg Arg Ala Ile Ser Met Ala Ala Lys Leu 995 1000 1005

Glu Ala Glu Val Arg Ala Arg Glu Arg Asn Thr Arg Met Gly Glu Pro 1010 1015 1020

Glu Gly His Glu Glu Pro Arg Gly Ala Glu Glu Ser Ile Ser Ala Leu 1025 1030 1035 1040

Gly Asp Leu Phe Ala Asp Leu Lys Phe Ala Leu Ser Glu Glu Asp Pro 1045 1050 1055

Trp Lys Ala Phe Glu Phe Leu Lys His Ala Trp Lys Ile Ala Gly Lys
1060 1065 1070

Ile Arg Leu Lys Pro Thr Cys Ser Phe 1075 1080

<210> 20 <211> 24 <212> DNA

<213> Artificial sequence

WO 99/19492 PCT/EP98/06977

18

<220>		
<223>	MSH6 specific primer 638 for PCR using cDNA of Arabidopsi thaliana ecotype Columbia	.s
<400>	20	
tctctaccag	g grgacgaaaa accg	24
<210>	21	
<211>	28	
<212>	DNA	
<213>	Artificial sequence	
<220>		
<223>	Primer S81 for PCR using cDNA of Arabidopsis thaliana eco Columbia	cype
<400>	21	
cgtcgccttt	agcatcccct tccttcac	₽8
	a.	
<210>	22	
<211>	30	
<212>	DNA	
<213>	Artificial sequence	
<220>	· ·	
<223>	MSH6 specific primer S823 for PCR using cDNA of Arabidops thaliana ecotype Columbia	sis
<400>	22	
gcttggcgca	a totaatagaa toatgacagg	30
<210>	23	
<211>	24	
<212>	DNA	
<213>	Artificial sequence	
<220>		
<223>	MSH6 specific primer 637 for PCR using cDNA of Arabidops thaliana ecotype Columbia	is
<400>	23	
gacagcgtca	a gttcttcaga atgc	24
<210>	24	
<211>	33	
<212>	DNA	

<213>	Artificial sequence	
<220 > < 223 >	MSH6 specific primer 1S8 for PCR using cDNA of Arabidops:	is
<400>	24	
atcccgggat	gcagegecag agategattt tgt	3:
<210>	25	
<211>	27	
<212>	DNA	
<213>	Artificial sequence	
<220>		
<223>	MSH6 specific primer S83 for PCR using cDNA of Arabidopsi thaliana ecotype Columbia	.s
<400>	25	
cgctatctat	ggctgcttcg aattgag	27
<210>	26	
<211>	1385	
<212>	DNA	
<213>	Arabidopsis thaliana ecotype Columbia	
<223>	Clone 43	
<400>	26	
cccgggatgc	agegeeagag aregatitig tettiettee aaaaaceeae ggeggegaet	60
acgaagggtt	tggtttccgg cgatgctgct agcggcgggg gcggcagcgg aggaccacga	120
tttaatgtga	aggaagggga tgctaaaggc gacgcttctg tacgttttgc tgtttcgaaa	180
tctgtcgatg	aggttagagg aacggatact ccaccggaga aggttccgcg tcgtgtcctg	240
ccgtctggat	ttaageegge tgaateegee ggtgatgett egteeetgtt eteeaatatt	300
atgcataagt	ttgtaaaagt cgatgatcga gattgttctg gagagaggag ccgagaagat	360
gttgttccgc	tgaatgattc atctctatgt atgaaggcta atgatgttat tcctcaattt	420
cgttccaata	atggtaaaac tcaagaaaga aaccatgctt ttagtttcag tgggagagct	480
gaacttagat	cagtagaaga tataggagta gatggcgatg ttcctggtcc agaaacacca	540
gggatgcgtc	cacgtgette tegettgaag egagttetgg aggatgaaat gaettttaag	600
gaggataagg	ttcctgtatt ggactctaac aaaaggctga aaatgctcca ggatccggtt	660

WO 99/19492

cgcggagaga	agaaagaagt	aaacgaagga	accaaatttg	aatggcttga	gtcttctcga	720
atcagggatg	ccaatagaag	acgtcctgat	gatccccttt	acgatagaaa	gaccttacac	780
ataccacctg	atgttttcaa	gaaaatgtct	gcatcacaaa	agcaatattg	gagtgttaag	840
agtgaatata	tggacattgt	gcttttcttt	aaagtgggga	aattttatga	gctgtatgag	900
ctagatgcgg	aattaggtca	caaggagctt	gactggaaga	tgaccatgag	tggtgtggga	960
aaatgcagac	aggttggtat	ctctgaaagt	gggatagatg	aggcagtgca	aaagctatta	1020
gctcgtggat	ataaagttgg	acgaatcgag	cagctagaaa	catctgacca	agcaaaagcc	1080
agaggtgcta	atactataat	tccaaggaag	ctagttcagg	tattaactcc	atcaacagca	1140
agcgagggaa	acatcgggcc	tgatgccgtc	catcttcttg	ctataaaaga	gatcaaaatg	1200
gagctacaaa	agtgttcaac	tgtgtatgga	tttgcttttg	ttgactgtgc	tgccttgagg	1250
ttttgggttg	ggtccatcag	cgatgatgca	tcatgtgctg	ctcttggagc	gttattgatg	1320
caggiticic	caaaggaagt	gttatatgac	agtaaagggc	tatcaagaga	agcacaaaag	1330
gctctaagga	aatatacgtt	gacagggtct	acggcggtac	agttggctcc	agtaccacaa	1440
gtaatggggg	atacagatgc	tgctggagtt	agaaatataa	tagaatctaa	cggatacttt	1500
aaaggttctt	ctgaatcatg	gaactgtgct	gttgatggtc	taaatgaatg	tgatgttgcc	1560
cttagtgctc	ttggagagct	aattaatcat	ctgtctaggc	taaagctaga	agatgtactt	1620
aagcatgggg	atattttcc	ataccaagtt	tacaggggtt	gtctcagaat	tgatggccag	1680
acgatggtaa	atcttgagat	atttaacaat	agctgtgatg	gtggtccttc	agggaccttg	1740
tacaaatatc	ttgataactg	tgttagtcca	actggtaagc	gactcttaag	gaattggatc	1800
tgccatccac	tcaaagatgt	agaaagcatc	aataaacggc	ttgatgtagt	tgaagaattc	1860
acggcaaact	cagaaagtat	gcaaatcact	ggccagtatc	tccacaaact	tccagactta	1920
gaaagactgc	tcggacgcat	caagtctagc	gttcgatcat	cagcctctgt	gttgcctgct	1980
cttctgggga	aaaaagtgct	gaaacaacga	gttaaagcat	ttgggcaaat	tgtgaaaggg	2040
ttcagaagtg	gaattgatct	gttgttggct	ctacagaagg	aatcaaatat	gatgagtttg	2100
ctttataaac	tctgtaaact	tcctatatta	gtaggaaaaa	gcgggctaga	gttatttctt	2160
tctcaattcg	aagcagccat	agatagcg				2188

<210 > <211 > <212 > <212 > <213 > <223 >	Clone 62	osis chaliar	ia ecotype (Columbia		
<400>	27					
catcagccto	tgtgttgcct	getettetgg	g ggaaaaaagt	gctgaaacaa	cgagttaaag	60
catttgggca	aattgtgaaa	gggttcagaa	a gtggaattga	tctgttgttg	gctctacaga	120
aggaatcaaa	tatgatgagt	ttgctttata	a actctgtaa	acttcctata	ttagtaggaa	180
aaagcgggct	agagttattt	ctttctcaat	tcgaagcagc	: catagatago	gactttccaa	240
attatcagaa	ccaagatgtg	acagatgaaa	acgetgaaac	: tctcacaata	cttatcgaac	300
tttttatcga	aagagcaact	caatggtctg	aggtcattca	caccataago	tgcctagatg	360
tcctgagatc	ttttgcaatc	gcagcaagto	tctctgctgg	aagcatggco	aggcctgtta	420
tttttcccga	atcagaagct	acagatcaga	atcagaaaac	aaaagggcca	atacttaaaa	480
tccaaggact	atggcatcca	tttgcagttg	cagccgatgg	tcaattgcct	gttccgaatg	540
atatactcct	tggcgaggct	agaagaagca	gtggcagcat	tcatcctcgg	tcattgttac	600
tgacgggacc	aaacatgggc	ggaaaatcaa	ctcttcttcg	tgcaacatgt	ctggccgtta	660
tctttgccca	acttggctgc	tacgtgccgt	gtgagtcttg	cgaaatctcc	ctcgtggata	720
ctatcttcac	aaggcttggc	gcatctgata	gaatcatgac	aggagagagt	acctttttgg	780
tagaatgcac	tgagacagcg	tcagttcttc	agaatgcaac	tcaggattca	ctagtaatcc	840
ttgacgaact	gggcagagga	actagtactt	tcgatggata	cgccattgca	tactcggttt	900
ttcgtcacct	ggtagagaaa	gttcaatgtc	ggatgctctt	tgcaacacat	taccaccctc	960
	•			acacatggct		1020
				cctagtgttc		1080
				ggcactcatg		1140
caaaccaagt						1200
aaaacttcaa						1260
agtcattggt		•				1320
acgacacttt						1320

ccggg	1	1385
<210>	28	
<211>	34	
<212>	DNA	
<213>	Artificial sequence	
<220>		
<223>	MSH6 specific primer 2S8 for PCR using cDNA of Arabidopsi thaliana ecotype Columbia	.s
<400>	26	
atcccgggtt	atttgggaac acagtaagag gatt	34
-210-	· ·	
<210><211>	29 27	
<211>	DNA	
	Artificial sequence	,
<213>	Arcilicial sequence	•
<220>		
<223>	MSH6 specific primer S82 for PCR using cDNA of Arabidopsi thaliana ecotype Columbia	
<400>	29	
gcgttcgatc	atcagcctct gtgttgc	. , 2
<210>	30	
<211>	3606	
<212>	DNA	
<213>	Arabidopsis thaliana ecotype Columbia	
<220>		
<221>	CDS	
<222>	(142)(3468)	
<223>	AtMSH6 full-length cDNA and deduced sequence of the encompolypeptide	ied
<400>	30	
aaaagttgag	ceetgaggag tategtttee gecattteta egacgeaagg egaaaatttt	60
tggcgccaat	ctttccccc tttcgaattc tctcagctca aaacatcgtt tctctctcac	120
tctctctcac	aattccaaaa a atg cag cgc cag aga tcg att ttg tct ttc Met Gln Arg Gln Arg Ser Ile Leu Ser Phe	171

ttc Phe	caa Gln	aaa Lys	ccc Pro	acc Thr 15	gcg Ala	gcg Ala	act Thr	acg Thr	aag Lys 20	ggt Gly	ttg Leu	gtt Val	tcc Ser	ggc Gly 25	gat Asp	219
gct Ala	gct Ala	agc Ser	ggc 30	GJA aaa	ggc	ggc Gly	agc Ser	gga Gly 35	gga Gly	cca Pro	cga Arg	ttt Phe	aat Asn 40	gtg Val	aag Arg	267
gaa Glu	ej aaa	gat Asp 45	gct Ala	aaa Lys	ggc Gly	gac Asp	gct Ala 50	ser	gta Val	cgt Arg	ttt Phe	gct Ala 55	gtt Val	tcg Ser	aaa Lys	315
tct Ser	gtc Val 60	gat Asp	gag Glu	gtt Val	aga Arg	gga Gly 65	acg Thr	gat Asp	act Thr	cca Pro	ccg Pro 70	gag Glu	aag Lys	gtt Val	ccg Pro	363
cgt Arg 75	cgt Arg	gtc. Val	ctg Leu	ccg Pro	tct Ser 80	gga Gly	ttt Phe	aag Lys	ccg Pro	gct Ala 85	gaa Glu	tcc Ser	gcc Ala	gst Gly	gat Asp 90	411
gct Ala	tcg Ser	tcc Ser	ctg Leu	ttc Phe 95	tcc Ser	aat Asn	att	atg Met	cat His 100	aag Lys	ttt Phe	gta Val	aaa Lys	gtc Val 105	gat Asp	459
gat Asp	cga Arg	gat Asp	tgt Cys 110	tct Ser	gga Gly	gag Glu	agg Arg	agc Ser 115	cga Arg	gaa Glu	gat Asp	gtt Val	gtt Val 120	ccg Pro	ctg Leu	507
aat Asn	gat Asp	tca Ser 125	tct Ser	cta Leu	tgt Cys	atg Met	aag Lys 130	gct Ala	aat Asn	gat Aşp	gtt Val	att Ile 135	cct Pro	caa Gln	ttt Phe	555
cgt Arg	ser 140	aat Asn	aat Asn	ggt	aaa Lys	act Thr 145	caa Gln	gaa Glu	aga Arg	aac Asn	cat His 150	gct Ala	ttt Phe	agt Ser	ttc Phe	603
agt Ser 155	GJÀ aaa	aga Arg	gct Ala	gaa Glu	ctt Leu 160	aga Arg	tca Ser	gta Val	gaa Glu	gat Asp 165	ata Ile	gga Gly	gta Val	gat Asp	ggc Gly 170	651
gat Asp	gtt Val	cct Pro	ggt Gly	cca Pro 175	gaa Glu	aca Thr	cca Pro	gly ggg	atg Met 180	cgt Arg	cca Pro	cgt Arg	gct Ala	tct Ser 185	cgc Arg	699
ttg Leu	aag Lys	cga Arg	gtt Val 190	ctg Leu	gag Glu	gat Asp	gaa Glu	atg Met 195	act Thr	ttt Phe	aag Lys	gag Glu	gat Asp 200	aag Lys	gtt Val	747
cct Pro	gta Val	ttg Leu 205	gac Asp	tct Ser	aac Asn	aaa Lys	agg Arg	ctg Leu	aaa Lys	atg Met	ctc Leu	cag Gln 215	gat Asp	ccg Pro	gtt Val	795

Cys												ttt Phe					843
gag Glu 235			-			_				_	_						891
												gtt Val					939
												agt Ser					987
												gag Glu 295					1035
												aag Lys					1083
												gaa Glu			ata Ile 330	ý.	1131
															cga Arg		1179
									Ala			aga Arģ			aat Asn		1227
act Thr	ata Ile	att Ile 365	Pro	agg Arg	aag Lys	cta Leu	gtt Val 370	Gln	gta Val	tta Leu	act Thr	cca Pro 375	Ser	aca Thr	gca Ala		1275
agc Ser	gag Glu 380	Gly	aac Asn	atc Ile	ggg Gly	cct Pro 385	Asp	gcc Ala	gto Val	cat His	Leu 390	Leu	gct Ala	ata Ile	aaa Lys		1323
g ag Glu 395	atc	aaa Lys	atg Met	gag Glu	Cta Leu 400	Gln	aag Lys	tgt Cys	tca Ser	act Thr 405	· Val	tat Tyr	gga Gly	ttt Phe	gct Ala 410		1371
ttt Phe	gtt Val	gac Asp	tgt Cys	get Ala 415	Ala	ttg Leu	agg Arg	ttt Phe	tgg Trp 420	Val	Gly	g tcc / Ser	ato : Ile	ago Ser 425	gat Asp		1419

					gct Ala										14	167
					gac Asp										15	515
					acg Thr										15	563
					atg Met 480			_	-	-		_	_		16	611
		_			gga Gly						_				10	659
-	_	_	_		cta Leu		-	_	-	_		_	_		1	707
					cat His				-		-	_	_		1	755
					ttt Phe						_				1:	803
	-		_	_	atg Met 560	_							-	_	1	851
					ggg										1	899
_				_	cga 'Arg						_				1	947
					atc Ile										1	995
					agt Ser						Tyr				2	043

					aga Arg 640											2091
tca Ser	tca Ser	gcc Ala	tct Ser	gtg Val 655	ttg Leu	cct Pro	gct Ala	ctt Leu	ctg Leu 660	GJA aaa	aaa Lys	aaa Lys	gtg Val	ctg Leu 665	aaa Lys	2139
					ttt Phe											2187
	-	_	_	_	gct Ala		_	_	_			_	_	_	_	2235
				_	aaa Lys					_			_			2283
					caa Gln 720											233 Î
		_			gat Asp	-		_	-		-	_				2379
					ttt Phe											2427
				_	tgc Cys		-	_	_	_			_		_	2475
					gga Gly										gaa Glu	2523
					cag Gln 800											2571
					cat His					Ala						2619
					ata Ile				Glu					Ser		2667

agc Ser	att Ile	cat His 845	cct Pro	cgg Arg	tca Ser	ttg Leu	tta Leu 950	ctg Leu	acg Thr	gga Gly	cca Pro	aac Asn 855	atg Met	ggc	gga Gly	7	2715
aaa Lys	ser 860	act Thr	ctt Leu	ctt Leu	cgt Arg	gca Ala 865	aca Thr	tgt Cys	ctg Leu	gcc Ala	gtt Val 870	atc Ile	ttt Phe	gcc Ala	caa Gln	7	2763
ctt Leu 875	ggc	tgc Cys	tac Tyr	gtg Val	ccg Pro 880	tgt Cys	gag Glu	tct Ser	tgc Cys	gaa Glu 885	atc Ile	tcc Ser	ctc Leu	gtg Val	gat Asp 890	2	2811
act Thr	atc Ile	ttc Phe	aca Thr	agg Arg 895	ctt Leu	ggc Gly	gca Ala	tct Ser	gat Asp 900	aga Arg	atc Ile	atg Met	aca Thr	gga Gly 905	gag Glu	2	2859
agt Ser	acc Thr	ttt Phe	ttg Leu 910	gta Val	gaa Glu	tgc Cys	act Thr	gag Glu 915	aca Thr	gcg Ala	tca Ser	gtt Val	ctt Leu 920	cag Gln	aat Asn	2	2907
gca Ala	act Thr	cag Gln 925	gat Asp	tca Ser	cta Leu	gta Val	atc Ile 930	ctt Leu	gac Asp	gaa Glu	ctg Leu	ggc Gly 935	aga Arg	gga Gly	act Thr	2	2955
agt Ser	act Thr 940	ttc Phe	gat Asp	gga Gly	tac Tyr	gcc Ala 945	acc Ile	gca Ala	tac Tyr	tcg Ser	gtt Val 950	ttt Phe	cgt Arg	cac His	ctg Leu	-	3003
gta Val 955	gag Glu	aaa Lys	gtt Val	caa Gln	tgt Cys 960	cgg Arg	atg Met	ctc Leu	ttt Phe	gca Ala 965	aca Thr	cat His	tac Tyr	cac	cct Pro 970	3	3051
ctc Leu	acc Thr	aag Lys	gaa Glu	ttc Phe 975	gcg Ala	tct Ser	cac His	cca Pro	cgt Arg 980	gtc Val	acc Thr	ccg Ser	aaa Lys	cac His 985	atg Met	3	3099
gct Ala	tgc Cys	gca Ala	ttc Phe 990	aaa Lys	tca Ser	aga Arg	tct Ser	gat Asp 995	tat Tyr	caa Gln	cca Pro	Arg	ggt Gly L000	tgt Cys	gat Asp	3	3147
caa Gln	Asp.	cta Leu 1005	gtg Val	ttc Phe	ttg Leu	Tyr	cgt Arg 1010	tta Leu	acc Thr	gag Glu	Gly	gct Ala 1015	tgt Cys	cct Pro	gag Glu	3	3195
Ser	tac Tyr 1020	gga Gly	ctt Leu	caa Gln	Val	gca Ala L025	ctc Leu	atg Met	gct Ala	Gly	ata Ile 1030	cca Pro	aac Asn	caa Gln	gtg Val	3	3243
gtt Val 1039	Glu	aca Thr	gca Ala	Ser	ggt Gly L040	gct Ala	gct Ala	caa Gln	Ala	atg Met 1045	aag Lys	aga Arg	tca Ser	Ile	020 GJA aaa	3	3291

gga aac tto Glu Asn Phe	aag tca Lys Ser 1055	agt gag Ser Glu	Leu Arg	tct gag Ser Glu 1060	ttc tca Phe Ser	agt ctg Ser Leu 1065	cat I	3339
gaa gac tgg Glu Asp Trp	teu Lys 1070	tca ttg Ser Leu	gtg ggt Val Gly 1075	att tct Ile Ser	Arg Val	gcc cac Ala His 1080	aac I	3387
aat gcc ccc Asn Ala Pro 1085	o Ile Gly	Glu Asp	gac tac Asp Tyr 1090	gac act Asp Thr	ttg ttt Leu Phe 1095	tgc tta Cys Leu	tgg 3 Trp	3435
cat gag ato His Glu Ile 1100					taaatgg	cta	3	3478
tgacataaca	ctatctgaa	g ctcgti	taagt ct	tttgcctc	tctgatg	ttt attc	ctctta 3	538
aaaaatgctt	atatatcaa	a aaatt	gttte et	cgattaaa	aaaaaaa	aaa aaaa	aaaaaa 3	598
aaaaaaa		,					3	60\$
.210					•			٠.
<210> <211>	31 1109							
<212> <213>	PRT Arabido	neie th	aliana e	cotype Co	alumbia			4. :
	AT UDIC	para cm						
<223>	Polypep	tide MSF		cotype co	JI GII WIA			
<223> <400>	Polypep 31	otide MSF		colipc co	orumbia.			
	31		46 ·			Pro Thr	Ala	
<400> Met Gln Arg	31 Gln Arg 5	Ser Ile	i6 Leu Ser	Phe Phe	Gln Lys	15		
<400> Met Gln Arg	31 Gln Arg 5 Lys Gly 20 Gly Pro	Ser Ile Leu Val	Leu Ser Ser Gly 25	Phe Phe 10 Asp Ala	Gln Lys Ala Ser	Gly Gly 30	Gly	
<400> Met Gln Arg 1 Ala Thr Thr	Gln Arg 5 Lys Gly 20 Gly Pro	Ser Ile Leu Val Arg Phe	Leu Ser Ser Gly 25 Asn Val	Phe Phe 10 Asp Ala Arg Glu	Gln Lys Ala Ser Gly Asp 45	Gly Gly 30 Ala Lys	Gly	
<400> Met Gln Arg 1 Ala Thr Thr Gly Ser Gly 35 Asp Ala Ser	31 Gln Arg 5 Lys Gly 20 Gly Pro Val Arg	Ser Ile Leu Val Arg Phe Phe Ala 55	Leu Ser Ser Gly 25 Asn Val 40 Val Ser	Phe Phe 10 Asp Ala Arg Glu Lys Ser	Gln Lys Ala Ser Gly Asp 45 Val Asp 60	Gly Gly 30 Ala Lys Glu Val	Gly Gly Arg	
<400> Met Gln Arg 1 Ala Thr Thr Gly Ser Gly 35 Asp Ala Ser 50 Gly Thr Asp	31 Gln Arg 5 Lys Gly 20 Gly Pro Val Arg Thr Pro	Ser Ile Leu Val Arg Phe Phe Ala 55 Pro Glu 70	Leu Ser Ser Gly 25 Asn Val 40 Val Ser Lys Val	Phe Phe 10 Asp Ala Arg Glu Lys Ser Pro Arg 75	Gln Lys Ala Ser Gly Asp 45 Val Asp 60 Arg Val	Gly Gly 30 Ala Lys Glu Val Leu Pro	Gly Gly Arg Ser 80	
<400> Met Gln Arg 1 Ala Thr Thr Gly Ser Gly 35 Asp Ala Ser 50 Gly Thr Asp 65	Gln Arg S Lys Gly 20 Gly Pro Val Arg Thr Pro Pro Ala 85 His Lys 100	Ser Ile Leu Val Arg Phe Phe Ala 55 Pro Glu 70 Glu Ser Phe Val	Leu Ser Ser Gly 25 Asn Val 40 Val Ser Lys Val Ala Gly Lys Val 105	Phe Phe 10 Asp Ala Arg Glu Lys Ser Pro Arg 75 Asp Ala 90 Asp Asp	Gln Lys Ala Ser Gly Asp 45 Val Asp 60 Arg Val Ser Ser Arg Asp	Gly Gly 30 Ala Lys Glu Val Leu Pro Leu Phe 95 Cys Ser 110	Gly Gly Arg Ser 80 Ser	

Me		/s 30	Ala	Asn	Asp	Val	Ile 135	Pro	Gln	Phe	Arg	Ser 140	Asn	Asn	Gly	Lys
Th		ln	Glu	Arg	Asn	His 150	Ala	Phe	Ser	Phe	Ser 155	Gly	Arg	Ala	Glu	Leu 160
Ar	g Se	er	Val	Glu	Asp 165	Ile	Gly	Val	Asp	Gly 170	Asp	Val	Pro	Gly	Pro 175	Glu
Th	r Pi	ro	Gly	Met 180	Arg	Pro	Arg	Ala	Ser 185	Arg	Leu	Lys	Arg	Val 190	Leu	Glu
As	p G		Met 195	Thr	Phe	Lys	Glu	Asp 200	Lys	Val	Pro	Val	Leu 205	Asp	Ser	Asn
Ly		rg 10	Leu	Lys	Met	Leu	Gln 215	Asp	Pro	Val	Cys	Gly 220	Glu	Lys	Lys	Glu
Va 22		sn	Glu	Gly	Thr	Lys 230	Phe	Glu	Trp	Leu	Glu 235	Ser	Ser	Arg	Ile	Arg 240
As	p A	la	Asn	Arg	Arg 245	Arg	Pro	Asp	Asp	Pro 250	Leu	Tyr	Asp	Arg	Lys 255	Thr
Le	u H	is	Ile	Pro 260	Pro	Asp	Val	Phe	Lys 265	Lys	Met	Ser	Ala	Ser 270	Gln	Lys
Gl	n T		Trp 275	Ser	Val	Lys	Ser	Glu 280	Tyr	Met	Asp	Ile	Val 285	Leu	Phe	Phe
Ly		al 90	Gly	Lys	Phe	Tyr	Glu 295	Leu	Tyr	Glu	Leu	Asp 300	Ala	Glu	Leu	Gly
Hi 30		ys	Glu	Leu	_	Trp 310	Lys	Met	Thr	Met	Ser 315	Gly	Val	Gly	Lys	Cys 320
Ar	g G	ln	Val	Gly	Ile 325	Ser	Glu	Ser	-	Ile 330	-	Glu	Ala	Val	Gln 335	-
Le	u Le	eu	Ala	Arg 340	Gly	Tyr	Lys	Val	Gly 345	Arg	Ile	Glu	Gln	Leu 350	Glu	Thr
Se	r A	gp	Gln 355	Ala	Lys	Ala	Arg	Gly 360	Ala	Asn	Thr	Ile	Ile 365	Pro	Arg	Lys
Le		al 70	Gln	Val	Leu	Thr	Pro 375	Ser	Thr	Ala	Ser	Glu 380	Gly	Asn	Ile	Gly
Pr 38		sp	Ala	Val	His	Leu 390	Leu	Ala	Ile	Lys	Glu 395	Ile	Lys	Met	Glu	Leu 400
Gl	n L	ys	Cys	Ser	Thr							Val	_	Cys	Ala 415	

Leu	Arg	Phe	Trp 420	Val	Gly	Ser	Ile	Ser 425	Asp	Asp	Ala	Ser	Cys 430	Ala	Ala
Leu	Gly	Ala 435	Leu	Leu	Met	Gln	Val 440	Ser	Pro	Lys	Glu	Val 445		туг	Asp
Ser	Lys 450	Gly	Leu	Ser	Arg	Glu 455	Ala	Gln	Lys	Ala	Leu 460	Arg	Lys	туг	Thr
Leu 465	Thr	Gly	Ser	Thr	Ala 470	Val	Gln	Leu	Ala	Pro 475	Val	Pro	Gln	Val	Met 480
Gly	Asp	Thr	Asp	Ala 485	Ala	Gly	Val	Arg	Asn 490	Ile	Ile	Glu	Ser	Asn 495	_
			500				Ser	505					510		
Asn	Glu	Cys 515	Asp	Val	Ala	Leu	Ser 520	Ala	Leu	Gly	Glu	Leu 525	Ile	Asn	His
	530					535	Asp		÷		540				
545					550		Cys			555					560
				565			Asn		570					575	_
Thr	Leu	Tyr	Lys 580	Tyr	Leu	Asp	Asn	Cys 585	Val	Ser	Pro	Thr	Gly 590	Lys	Arg
		595		-			His 600					605		•	
Asn	Lys 610	Arg	Leu	Asp	Val	Val 615	Glu	Glu	Phe	Thr	Ala 620	Asn	Ser	Glu	Ser
Met 625	Gln	Ile	Thr	Gly	Gln 630	Tyr	Leu	His	Lys	Leu 635	Pro	Asp	Leu	Glu	Arg 640
Leu	Leu	Gly	Arg	Ile 645	Lys	Ser	Ser	Val	Arg 650	Ser	Ser	Ala	Ser	Val 655	Leu
Pro	Ala	Leu	Leu 660	Gly	Lys	Lys	Val	Leu 665	Lys	Gln	Arg	Val	Lys 670	Ala	Phe
Gly	Gln	Ile 675	Val	Lys	Gly	Phe	Arg 680	Ser		Ile	Asp	Leu 685	Leu	Leu	Ala
Leu	Gln 690	Lys	Glu	Ser	Asn	Met 695	Met	Ser	Leu	Leu	Tyr	Lys	Leu	Cys	Lys

Leu 705	Pro	He	Leu	Val	Gly 710	Lys	Ser	Gly	Leu	Glu 715	Leu	Phe	Leu	Ser	Gln 720
Phe	Glu	Ala	Ala	Ile 725	Asp	Ser	Asp	Phe	Pro 730	Asn	Tyr	Gln	Asn	Gln 735	Asp
Val	Thr	Asp	Glu 740	Asn	Ala	Glu	Thr	Leu 745	Thr	Ile	Leu	Ile	Glu 750	Leu	Phe
Ile	Glu	Arg 755	Ala	Thr	Gln	Trp	Ser 760	Glu	Val	Ile	His	Thr 765	Ile	Ser	Суѕ
Leu	Asp 770	Val	Leu	Arg	Ser	Phe 775	Ala	Ile	Ala	Ala	Ser 780	Leu	Ser	Ala	Gly
Ser 785	Met	Ala	Arg	Pro	Val 790	Ile	Phe	Pro	Glu	Ser 795	Glu	Ala	Thr	Asp	Gln 800
Asn	Gln	Lys	Thr	Lys 805	Gly	Pro	Ile	Leu	Lys 810	Ile	Gln	Gly	Leu	Trp 815	His
			820					825					Asn 830	_	
Leu	Leu	Gly 835	Glu	Ala	Arg	Arg	Ser 840	Ser	Gly	Ser	Ile	His 845	Pro	Arg	Ser
Leu	Leu 850	Leu	Thr	Gly	Pro	Asn 855	Met	Ġly	Gly	Lys	Ser 860	Thr	Leu	Leu	Arg
Ala 865	Thr	Cys	Leu	Ala	Val 870	Ile	Phe	Ala	Gln	Leu 875	Gly	Cys	Tyr	Val	Pro 880
Cys	Glu	Ser	Cys	Glu 885	Ile	Ser	Leu	Val	Asp 890	Thr	Ile	Phe	Thr	Arg 895	Leu
Gly	Ala	Ser	Asp 900	Arg	Ile	Met	Thr	Gly 905	Glu	Ser	Thr	Phe	Leu 910	Val	Glu
Cys	Thr	Glu 915	Thr	Ala	Ser	Val	Leu 920	Gln	Asn	Ala	Thr	Gln 925	Asp	Ser	Leu
Val	Ile 930	Leu	Asp	Glu	Leu	Gly 935	Arg	Gly	Thr	Ser	Thr 940	Phe	Asp	Gly	Tyr
Ala 945	Ile	Ala	Tyr	Ser	Val 950	Phe	Arg	His	Leu	Val 955	Glu	Lys	Val	Gln	Cys 960
Arg	Met	Leu	Phe	Ala 965	Thr	His	Tyr	His	Pro 970	Leu	Thr	Lys	Glu	Phe 975	Ala
Ser	His	Pro	Arg 980	Val	Thr	Ser	Lys	His 985	Met	Ala	Cys	Ala	Phe 990	Lys	Ser

Arg Ser Asp T 995	Tyr Gln Pro Arg Gly Cys Asp Gln Asp Leu Val Phe Leu 1000 1005	
Tyr Arg Leu T 1010	Thr Glu Gly Ala Cys Pro Glu Ser Tyr Gly Leu Gln Val 1015 1020	
Ala Leu Met A 1025	Ala Gly Ile Pro Asn Gln Val Val Glu Thr Ala Ser Gly 1030 1035 1040	
Ala Ala Gln A	Ala Met Lys Arg Ser Ile Gly Glu Asn Phe Lys Ser Ser 1045 1050 1055	
	Ser Glu Phe Ser Ser Leu His Glu Asp Trp Leu Lys Ser 1060 1065 1070	
Leu Val Gly 1	Ile Ser Arg Val Ala His Asn Asn Ala Pro Ile Gly Glu 1080 1085	
Asp Asp Tyr A	Asp Thr Leu Phe Cys Leu Trp His Glu Ile Lys Ser Ser 1095 1100	•
Tyr Cys Val : 1105	Pro Lys	
	•	
<210>	32	
<211>	24	
<212>	DNA	
<213>	Artificial sequence	
<220>		
<223>	Forward primer for PCR amplification of ATHGENEA microsatellite	
<400>	32	
accatgcata g	cttaaactt cttg	24
<210>	33	
<210> <211>	22	
<211>	DNA	
<212>	Artificial sequence	
<220>	Description of ATUCENER	
<223>	Reverse primer for PCR amplification of ATHGENEA microsatellite	
<400>	33	
acataaccac a	aataggggt gc .	22

2 to 3 a 1 a 2 a 2 a 2 a 2 a 2

```
<210>
              34
 <211>
              18
 <212>
              DNA
 <213>
              Artificial sequence
 <220>
              Forward primer DMCIN-A for PCR on genomic DNA of Arabidopsis
 <223>
              thaliana ssp. Landsberg erecta "Ler"
 <400>
              34
 gaagcgatat tgttcgtg
                                                                           18
 <210>
              35
 <211>
              18
 <212>
              DNA
 <213>
              Artificial sequence
 <220>
 <223>
              Reverse primer DMCIN-B for PCR on genomic DNA of Arabidopsis
              thaliana ssp. Landsberg erecta "Ler"
<400>
              35
agattgcgag aacattcc
                                                                           18
<210>
              36
<211>
              31
<212>
              DNA
<213>
              Artificial sequence
<220>
<223>
             Forward primer DMCIN-1 for PCR on genomic DNA of Arabidopsis
              thaliana ssp. Landsberg erecta "Ler"
<400>
             36
acgcgtcgac tcagctatga gattactcgt g
                                                                          31
<210>
             37
<211>
             29
<212>
             DNA
<213>
             Artificial sequence
<220>
<223>
             Reverse primer DMCIN-2 for PCR on genomic DNA of Arabidopsis
             thaliana ssp. Landsberg erecta "Ler"
<400>
             37
gctctagatt tctcgctcta agactctct
                                                                          29
```

<210>	38
<211>	32
<212>	DNA
<213>	Artificial sequence
<220>	
<223>	Forward primer DMCIN-3 for PCR on genomic DNA of Arabidopsis thaliana ssp. Landsberg erecta "Ler"
<400>	38
gctctagagc	ttctcttaag taagtgattg at 32
<210>	39
<211>	48
<212>	DNA
<213>	Artificial sequence
<220>	•
<223>	Reverse primer DMCIN-4 for PCR on genomic DNA of Arabidopsis thaliana ssp. Landsberg erecta "Ler"
<400>	39
tcccccgggc	togagagate tecatggttt etteagetet atgaatee 48
<210>	40
<211>	26
<212>	DNA
<213>	Artificial sequence
<220>	
<223>	Forward primer DMC1a for PCR on genomic DNA of Arabidopsis thaliana ssp. Landsberg erecta "Ler"
<400>	40
acgcgtcgac	gaattcgcaa gtgggg 26
<210>	41
<211>	38
<212>	DNA
<213>	Artificial sequence
<220>	
<223>	Reverse primer DMC1b for PCR on genomic DNA of Arabidopsis thaliana ssp. Landsberg erecta "Ler"
<400>	41

tccatggaga	tctcccgggt accgatttgc ttcgaggg 38
<210>	4.2
<211>	20
<212>	DNA
<213>	
(21)/	Artificial sequence
<220>	
<223>	Forward primer for PCR amplification of ATEAT1 SSLP marker in Arabidopsis thaliana subspecies
<400>	42
gccactgcgt	gaatgatatg 20
<210>	43
<211>	22
<212>	DNA
<213>	Artificial sequence
<220>	
<223>	Poytowno maimon for pop and list
(223)	Reverse primer for PCR amplification of ATEAT1 SSLP marker in Arabidopsis thaliana subspecies
<400>	43
cgaacagcca	acattaattc cc 22
<210>	44
<211>	18
<212>	DNA
<213>	Artificial sequence
	crricial sequence
<220>	•
<223>	Forward primer for PCR amplification of NGA63 SSLP marker in Arabidopsis thaliana subspecies
<400>	44
aaccaaggca	cagaagcg 18
<210>	45
<211>	18
<212>	DNA
<213>	Artificial sequence
:220>	
:223>	Reverse primer for PCR amplification of NGA63 SSLP marker in
	Arabidopsis thaliana subspecies

<400>	45	
acccaagtga	tegecace	18
<210>	46	
<211>	21	
<212>	DNA	
<213>	Artificial sequence	
	•	
<220>		
<223>	Forward primer for PCR amplification of NGA248 SSLP marker Arabidopsis thaliana subspecies	in
<400>	46	
taccgaacca	aaacacaaag g	21
-210-	47	
<210>	47	•
<211>	22	
<212>	DNA	
<213>	Artificial sequence	
<220>	At .	
<223>	Reverse primer for PCR amplification of NGA248 SSLP marker Arabidopsis thaliana subspecies	in
<400>	47	
tctgtatctc	ggtgaattet ee	22
<210>	48	
<211>	22	
<212>	DNA ·	
<213>	Artificial sequence	
<220>		
<223>	Forward primer for PCR amplification of NGA128 SSLP marker	
	Arabidopsis thaliana subspecies	
<400>	48	
ggtctgttga	tgtcgtaagt cg	22
<210>	49	
<211>	22	
<212>	DNA	
<213>	Artificial sequence	
<220>	p	

	·
<223>	Reverse primer for PCR amplification of NGA128 SSLP marker in Arabidopsis thaliana subspecies
<400>	49
atcttgaaac o	ctttagggag gg 22
<210>	50
<211>	22
<212>	DNA
<213>	
	Artificial sequence
<220>	
<223>	Forward primer for PCR amplification of NGA280 SSLP marker in Arabidopsis thaliana subspecies
<400>	50
ctgatctcac g	gacaatagt gc €2
<210>	51
<211>	20
<212>	DNA
<213>	Artificial sequence
<220>	
<223>	Reverse primer for PCR amplification of NGA280 SSLP marker in Arabidopsis thaliana subspecies
<400>	51
ggctccataa aa	aagtgcacc 20
<210>	52
<211>	21
	DNA
	Artificial sequence
<220>	
<223>	Forward primer for PCR amplification of NGA111 SSLP marker in Arabidopsis thaliana subspecies
<400>	52
ctccagttgg aa	gctaaagg g
<210>	53
<211>	21
	DNA
<213>	Artificial sequence

<220>		
<223>	Reverse primer for PCR amplification of NGA111 SSLP marke Arabidopsis thaliana subspecies	er in
< 400 >	53	
tgttttttag	gacaaatggc g	21
<210>	54	
<211>	20	
<212>	DNA	
<213>	Artificial sequence	
<220>		
<223>	Forward primer for PCR amplification of NGA168 SSLP marke Arabidopsis thaliana subspecies	er in
<400>	54	
ccttcacatc	caaaacccac	20
<210>	55	
<211>		
<211>	20	
<213>	DNA	
(213)	Artificial sequence	
<220>		
<223>	Reverse primer for PCR amplification of NGA168 SSLP marke Arabidopsis thaliana subspecies	r in
<400>	55	
gcacataccc	acaaccagaa	20
<210>	56	
<211>	20	
<212>	DNA	
<213>	Artificial sequence	
<220>		
<223>	Forward primar for DCD and life and a constant constant	
~4437	Forward primer for PCR amplification of NGA1126 SSLP mark in Arabidopsis thaliana subspecies	er
<400>	56	
cgctacgctt	ttcggtaaag	20

Carrier of the State

<210>	57
<211>	20
<212>	DNA
<213>	
	Artificial sequence
<220>	
<223>	Payonna maio
1000	Reverse primer for PCR amplification of NGA1126 SSLP marker in Arabidopsis thalians subgression
	in Arabidopsis thaliana subspecies
<400>	57
1.002	57
gcacagto	ca agtcacaacc
3	20
<210>	58
<211>	20
<212>	DNA
<213>	
\213 <i>></i>	Artificial sequence
<220>	
<223>	
(243)	Forward primer for PCR amplification of NGA361 SSLP marker in
	Arabidopsis thaliana subspecies
<400>	
(400)	58
aaagagaro	ja gaatttggac
	ga gaattiggac 20
<210>	59
<211>	23
<212>	DNA
<213>	
1237	Artificial sequence
<220>	
<223>	Payana minara
	Reverse primer for PCR amplification of NGA361 SSLP marker in
	Arabidopsis thaliana subspecies
<400>	59
	33
acatatcaa	atattaaagt agc
	23
<210>	60
<211>	18
<212>	
<213>	DNA
~~13 3	Artificial sequence
<220>	
<223>	
~4437	Forward primer for PCR amplification of NGA168 SSLP marker in
	Arabidopsis thaliana subspecies
-400-	
<400>	60

WO 99/19492 PCT/EP98/06977

tegtetactg	cactgeeg 18
<210>	61
<211>	22
<212>	DNA
<213>	Artificial sequence
	carretur sequence
<220>	
<223>	Reverse primer for PCR amplification of NGA168 SSLP marker in Arabidopsis thaliana subspecies
<400>	61
gaggacatgt	ataggagcct cg 22
<210>	62
<211>	20
<212>	DNA
<213>	Artificial sequence
<220>	
<223>	Forward primer for PCR amplification of AthBIO2 SSLP marker in Arabidopsis thaliana subspecies
<400>	62
tgacctcctc	ttccatggag 20
<210>	63
<211>	22
<212>	DNA
<213>	Artificial sequence
<220>	
<223>	Reverse primer for PCR amplification of AthBIO2 SSLP marker in Arabidopsis thaliana subspecies
<400>	63
ttaacagaaa	cccaaagctt tc 22
<210>	64
<211>	21
<212>	DNA
<213>	Artificial sequence
<220>	
<223>	Forward primer for DCD amplification of Athrotograp corp
	Forward primer for PCR amplification of AthUBIQUE SSLP marker in Arabidopsis thaliana subspecies

Later to the

```
<400>
                64
  aggcaaatgt ccatttcatt g
                                                                             21
  <210>
               65
  <211>
               20
  <212>
               DNA
  <213>
               Artificial sequence
  <220>
               Reverse primer for PCR amplification of AthUBIQUE SSLP marker
  <223>
               in Arabidopsis thaliana subspecies
  <400>
               65
 acgacatggc agatttctcc
                                                                            20
 <210>
               66
 <211>
               21
 <212>
               DNA
 <213>
              Artificial sequence
 <220>
              Forward primer for PCR amplification of NGA172 SSLP marker in
 <223>
              Arabidopsis thaliana subspecies
 <400>
              66
 agetgettee ttatagegte e
                                                                            21
 <210>
              67
 <211>
              19
 <212>
              DNA
<213>
              Artificial sequence
<220>
<223>
             Reverse primer for PCR amplification of NGA172 SSLP marker in
             Arabidopsis thaliana subspecies
<400>
             67
catccgaatg ccattgttc
                                                                           19
<210>
             68
<211>
             21
<212>
             DNA
<213>
             Artificial sequence
<220>
```

<223>	Forward primer for PCR amplification Arabidopsis thaliana subspecies	of	NGA126	SSLP	marker i
<400×	63		•		
gaaaaaacgc	tactttcgtg g				21
<210>	69				
<211>	22				
<212>	DNA				
<213>	Artificial sequence				
	•				
<220>					
<223>	Reverse primer for PCR amplification Arabidopsis thaliana subspecies	of	NGA126	SSLP	marker i
<400>	69				
caagagcaat	atcaagagca gc				22
					•
<210>	70				
<211><212>	20				
<212>	DNA				
(213)	Artificial sequence				
<220>					
<223>	Forward primer for PCR amplification Arabidopsis thaliana subspecies	of	NGA162	SSLP	marker i
<400>	70				
catgcaattt	gcatctgagg				20
<210>	71		,		
<211>	22				
<212>	DNA				
<213>	Artificial sequence				
<220>					
<223>	Reverse primer for PCR amplification	٥f	NCA162	CCID	marker i
1227	Arabidopsis thaliana subspecies	O.L	NGA162	2205	marker 1
<400>	71				
ctctgtcact	cttttcctct gg				22
•					
<210>	72				•
<211>	21				
<212>	DNA				
<213>	Artificial sequence				

```
<220>
   <223>
                Forward primer for PCR amplification of NGA6 SSLP marker in
                Arabidopsis thaliana subspecies
   <400>
                72
   tggatttctt cctctcttca c
                                                                             21
  <210>
                73
  <211>
                21
  <212>
                DNA
  <213>
                Artificial sequence
  <220>
  <223>
               Reverse primer for PCR amplification of NGA6 SSLP marker in
               Arabidopsis thaliana subspecies
  <400>
               73
 atggagaage ttacactgat c
                                                                            21
 <210>
               74
 <211>
               20
 <212>
               DNA
 <213>
               Artificial sequence
 <220>
 <223>
              Forward primer for PCR amplification of NGA12 SSLP marker in
              Arabidopsis thaliana subspecies
 <400>
              74
aatgttgtcc tcccctcctc
                                                                           20
<210>
              75
<211>
              22
<212>
              DNA
<213>
              Artificial sequence
<220>
<223>
             Reverse primer for PCR amplification of NGA12 SSLP marker in
             Arabidopsis thaliana subspecies
<400>
             75
tgatgctctc tgaaacaaga gc
                                                                          22
```

<210>	76
<211>	21
<212>	DNA
<213>	Artificial sequence
<220>	
<223>	Forward primer for PCR amplification of NGA8 SSLP marker in Arabidopsis thaliana subspecies
<400>	76
gagggcaaat	ctttatttcg g
<210>	77
<211>	22
<212>	DNA
<213>	Artificial sequence
<220>	
<223>	Reverse primer for PCR amplification of NGA8 SSLP marker in Arabidopsis thaliana subspecies
<400>	77
tggctttcgt	ttataaacat cc 22
<210>	78
<211>	21
<212>	DNA
<213>	Artificial sequence
<220>	
<223>	Forward primer for PCR amplification of NGA1107 SSLP marker in Arabidopsis thaliana subspecies
<400>	78
gcgaaaaaac a	aaaaaaatcc a 21
<210>	79
<211>	21
<212>	DNA
<213>	Artificial sequence
<220>	
<223>	Reverse primer for PCR amplification of NGA1107 SSLP marker in Arabidopsis thaliana subspecies
<400>	79

cgacgaa	tcg acagaattag g
	. 2
<210>	80
<211>	21
<212>	DNA
<213>	Artificial sequence
<220>	
<223>	Forward primer for PCR amplification of NGA225 SSLP marker i
	Arabidopsis thaliana subspecies
<400>	80
gaaatcca	aaa tcccagagag g
	21
<210>	81
<211>	22
<212>	DNA
<213>	Artificial sequence
<220>	·
<223>	Reverse primer for PCR amplification of NGA225 SSLP marker in Arabidopsis thaliana subspecies
<400>	81
tctccccad	et agttttgtgt cc 22
<210>	
<211>	82
<212>	19 DNA
<213>	Artificial sequence
	ocquence
220>	
:223>	Forward primer for PCR amplification of NGA249 SSLP marker in Arabidopsis thaliana subspecies
400>	82
accgtcaa	t ttcatcgcc
210>	
210> 211>	83
212>	22 DNA
213>	Artificial sequence
220>	
23>	Reverse primer for PCR amplification of NGA249 SSLP marker in
	Arabidopsis thaliana subspecies

<400>	83
ggatocotaa	ctgtaaaatc cc 22
<210>	84
<211>	22
<212>	DNA
<213>	Artificial sequence
<220>	
<223>	Forward primer for PCR amplification of CA72 SSLP marker in Arabidopsis thaliana subspecies
<400>	84
aatcccagta	accaaacaca ca 22
<210>	85
<211>	20
<212>	DNA
<213>	Artificial sequence
	crriciar sequence
<220>	
<223>	Reverse primer for PCR amplification of CA72 SSLP marker in Arabidopsis thaliana subspecies
<400>	85
cccagtctaa	ccacgaccac 20
<210>	86
<211>	20
<212>	DNA
<213>	Artificial sequence
(213)	Arciricial sequence
<220>	
<223>	Forward primer for PCR amplification of NGA151 SSLP marker in
	Arabidopsis thaliana subspecies
<400>	86
gttttgggaa	gttttgctgg 20
<210>	· 07
<211>	87
<211> <212>	24
<212> <213>	DNA
~6137	Artificial sequence
<220>	•
	· · · · · · · · · · · · · · · · · · ·

e in this case, in the

```
Reverse primer for PCR amplification of NGA151 SSLP marker in
  <223>
               Arabidopsis thaliana subspecies
  <400>
               87
  cagtctaaaa gcgagagtat gatg
                                                                            24
  <210>
               88
  <211>
               22
 <212>
               DNA
 <213>
               Artificial sequence
 <220>
              Forward primer for PCR amplification of NGA106 SSLP marker in
 <223>
              Arabidopsis thaliana subspecies
 <400>
              88
 gttatggagt ttctagggca cg
                                                                           22
 <210>
              89
 <211>
              20
 <212>
              DNA
 <213>
              Artificial sequence
 <220>
 <223>
              Reverse primer for PCR amplification of NGA106 SSLP marker in
              Arabidopsis thaliana subspecies
<400>
              89
tgccccattt tgttcttctc
                                                                          20
<210>
              90
<211>
              20.
<212>
             DNA
<213>
           Artificial sequence
<220>
             Forward primer for PCR amplification of NGA139 SSLP marker in
<223>
             Arabidopsis thaliana subspecies
<400>
             90
agagctacca gatccgatgg
                                                                         20
<210>
             91
<211>
             21
<212>
             DNA
<213>
          F Artificial sequence
```

48

<220>	
<223>	Reverse primer for PCR amplification of NGA139 SSLP marker in Arabidopsis thaliana subspecies
<400>	91
ggtttcgttt	cactatccag g
<210>	92
<211>	22
<212>	DNA _
<213>	Artificial sequence
<220>	
<223>	Forward primer for PCR amplification of NGA76 SSLP marker in
	Arabidopsis thaliana subspecies
<400>	92
ggagaaaatg	tcactctcca cc 22
<210>	93
<211>	20
<212>	DNA
<213>	Artificial sequence
<220>	
<223>	Reverse primer for PCR amplification of NGA76 SSLP marker in
	Arabidopsis thaliana subspecies
<400>	93
aggcatggga	gacatttacg 20
<210>	94
<211>	20
<212>	DNA
<213>	Artificial sequence
	*
<220>	
<223>	Forward primer for PCR amplification of ATHSO191 SSLP marker in Arabidopsis thaliana subspecies
<400>	94
ctccaccaat	Catocaaato

gctaacaata aaggtatacg gttttatgtc atcaatataa ctatataa	60
97	
Genomic DNA sequence of AtMSH6	
Arabidopsis thallana ecotype Columbia	
	•
	. 22
s atggtettga gg	22
97	
Reverse primer for PCR amplification of NGA129 SSLP ma Arabidopsis thaliana subspecies	rker in
Artificial sequence	•
DNA	
22	
97	
a ctaaagtgag gg	22
96	
Forward primer for PCR amplification of NGA129 SSLP ma Arabidopsis thaliana subspecies	arker in
	•
•	
96 22	
at ggagatggtc a	21
95	
in Arabidopsis thaliana subspecies	marker
Reverse primer for PCR amplification of ATHSO191 SCID	manican
=	
95	
	21 DNA Artificial sequence Reverse primer for PCR amplification of ATHSO191 SSLP in Arabidopsis thaliana subspecies 95 St ggagatggtc a 96 22 DNA Artificial sequence Forward primer for PCR amplification of NGA129 SSLP material and subspecies 96 a ctaaagtgag gg 97 22 DNA Artificial sequence Reverse primer for PCR amplification of NGA129 SSLP material and subspecies 97 31 atggtcttga gg 98 8062 DNA Arabidopsis thaliana ecotype Columbia Genomic DNA sequence of AtMSH6 97

aagaaatgaa	agatatatat	tgtttttca	tttatcaaac	aaaacaacaa	gactttttt	120
ttacttttta	cattggtcaa	caaaatacaa	gataaacgac	atogittaat	catttcccaa	180
ttttacccct	aagtttaaca	cctagaacct	tctccatctt	cgcaagcaca	gcctgattag	240
gaacagcttt	accattctca	tattcctgaa	ctacctgagt	cctctcattg	atctgtttcg	300
ccaaatccgc	ttgtgacatc	ttcttctcca	atctcgcttt	ctgtatcatc	aacctcacct	360
ctgctttcac	acgatccatc	gccgcaggct	ctgtttcttc	ttccagcttc	ttcgtgttaa	420
tcaccggaac	cgccgtagat	ttcccctttt	tgttcgaacc	ggcatcgaat	ttcttaaccg	480
tttgaaccgc	gacaccgttt	ctcagagctg	cgttaaccgc	tttcggatcg	cgtaggtctt	540
ggctcttttg	ttttgatttg	tggagaacta	ctggttccca	gtcttgtgtt	actgctcctg	600
ggtatetget	cggcatcgtc	gatgaattga	gagaaaggaa	caacgcgaaa	attttattaa	660
tctgagtttt	gaaattgaga	aacgatgaag	atgaagaatg	tigitgagag	gattgtgata	720
tttatatata	cgaagattgg	tttctggaga	attcgatcat	ctttttctcc	attttcgtct	780
ctggaacgtt	cttagagatg	attgacgacg	tgtcattatc	tgatttgcag	ttaaccaatg	840
ctttttgggt	tggattcgtg	gtacaccata	ttatccgatt	tggctcaatg	gttttatata	900
aatttggttt	tcggttcggt	tatgagttat	cattaaaatt	aagctaacca	aaaattttcg	960
taaaatttat	ttcggtttca	attcggatcc	cttacttcca	gaaccgaatt	attcgaaacc	1020
ggggttagcc	gaaccgaata	ccaatgcctg	attgactcgt	tggctagaaa	gatccaacgg	1080
tatacaataa	tagaacataa	atcggacggt	catcaaagcc	tcaaagagtg	aacagtcaac	1140
aaaaaagtt	gagccctgag	gagtatcgtt	tccgccattt	ctacgacgca	aggcgaaaat	1200
ttttggcgcc	aatctttccc	ccctttcgaa	ttctctcagc	tcaaaacatc	gtttctctct	1260
cactetetet	cacaattcca	aaaaatgcag	cgccagagat	cgattttgtc	tttcttccaa	1320
aaacccacgg	cggcgactac	gaagggtttg	gtttccggcg	atgctgctag	cggcggggc	1380
ggcagcggag	accacgattt	aatgtgaagg	aaggggatgc	taaaggcgac	gcttctgtac	1440
gttttgctgt	ttcgaaatct	gtcgatgagg	ttagaggaac	ggatactcca	ccggagaagg	1500
tccgcgtcg	tgtcctgccg	tctggattta	agccggctga	atccgccggt	gatgcttcgt	1560
cctgttctc	caatattatg	cataagtttg	taaaagtcga	tgatcgagat	tgttctggag	1620
agaggtacta	atcttcgatt	ctcttaattt	tgttatcttt	agctggaaga	agaagattcq	1680

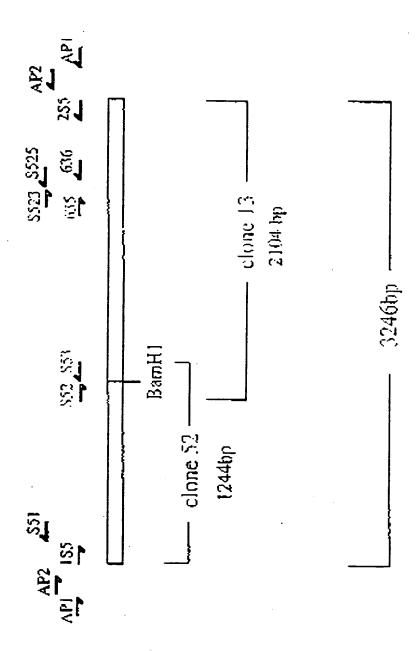
FORTERDANA	
tgtaatttgt tgtattcgtt ggagagattc tgattactgc attggatcgt tgtttacaaa	1740
ttttcaggag ccgagaagat gttgttccgc tgaatgattc atctctatgt atgaaggcta	1800
atgatgitat teeteaatti egiteeaata atggiaaaae teaagaaaga aaccatgett	1860
ttagtttcag tgggagagct gaacttagat cagtagaaga tataggagta gatggcgatg	1920
tteetggtee agaaacacca gggatgegte caegtgette tegettgaag egagttetgg	1980
aggatgaaat gacttttaag gaggataagg ttcctgtatt ggactctaac aaaaggctga	2040
aaatgeteea ggateeggtt tgtggagaga agaaagaagt aaaegaagga accaaatttg	2100
aatggettga gtettetega ateagggatg ecaatagaag aegteetgat gateceettt	
acgatagaaa gaccttacac ataccacctg atgttttcaa gaaaatgtct gcatcacaaa	2160
agcaatattg gagtgttaag agtgaatata tggacattgt gcttttcttt aaagtggtta	2220
gtaactatta atctagtgtt caatccattt cctcaatgtg atttgttcac ttacatctgt	2280
ttacgttatg ctcttctcag gggaaatttt atgagctgta tgagctagat gcggaattag	2340
gtcacaagga gcttgactgg aagatgacca tgagtggtgt gggaaaatgc agacaggtaa	2400
attagttgaa acaactggcc tgcttgaatt attgtgtcta taaattttga caccaccttt	2460
tgtttcaggt tggtatctct gaaagtggga tagatgaggc agtgcaaaag ctattagctc	2520
gtgggtaagg gaaccatcat actttatgga attcgtttac tgctacttcg gctaggattt	2580
aagaaatgga aatcacttca agcatcatta gttaggatcc tgagaactca ggatgttttc	2640
ttattcgtta tataataagt cttttcatca aggagtaaca aacaaaactt gcacaatatt	2700
tgtgtgctca ctggcaaggc atatataccc agctaacctt tgctagttca ctgtagtaac	2760
agttacggat aatatatgtt tacttgtatg tggtaccctc attttgtctc tcatggaggc	2820
tttcaagcot tgtgttgaaa ctggaraget accepte attttgtctc tcatggaggc	2880
tttcaagcct tgtgttgaaa ctggatagtt acatatgctt ccaacagaaa ctagcatgca gattcatatg ctttcctart gractaata	2940
gattcatatg ctttcctatt ctactaatta tgtattgaca cactcgttgt ttcttttgaa	3000
agatataaag ttggacgaat cgagcagcta gaaacatctg accaagcaaa agccagaggt	3060
gctaatactg taagttttct tggataggtc aaggagagtg ttgcagactg tttttgatca	3120
tttctttttc tgtacattac tttcatgctg taattaactc aatggctatt ctggtctgat	3180
tatcagataa ttecaaggaa getagtteag gtattaaete catcaacage aagegaggga	3240
aacatcgggc ctgatgccgt ccatcttctt gctataaaag aggtttgtta tttacttatt	3300

tatcttatca	tgttcagttc	atccaagtcc	tgaaaaatta	cactcttctt	taccaatctt	3360
ccatcaagct	gtgtaaagga	tttggaatta	gaaaatcatt	atttgatgct	ttgttttata	3420
tgcaagaggt	tcccttgaaa	agatctgttt	aagattcttt	gcacttgaaa	aattcaatct	3480
ttttaagtga	atcccctact	ttcttacaat	gatcatagtc	tgcaattgca	tgtcaagtaa	3540
tatcattcct	tgttactgca	tcccctctt	tcttaatgac	cattgtctat	gttgtgtttg	3600
tctcgtgtgc	tggagaaaat	gatagctgat	ccaagctgta	cattatcatg	attaagtagc	3660
tgctcaggaa	ttgcctttgg	ttacattgcc	taatggtttg	atgtcaattt	ttcttctgaa	3720
tctttattt	agatcaaaat	ggagctacaa	aagtgttcaa	ctgtgtatgg	atttgctttt	3780
gttgactgtg	ctgccttgag	gttttgggtt	gggtccatca	gcgatgatgc	atcatgtgct	3840
gctcttggag	cgttattgat	gcaggtaagc	aagtgtattc	tgtatcttat	gtgtaccatg	3900
tgacttcctg	tgcatatatt	tgggttgcag	gaactaattc	tgaatcacca	tttggtatgt	3960
tttttccagg	tttctccaaa	ggaagtgtta	tatgacagta	aaggtaaact	gcttgtatcg	4020
ccagttgttt	tgttaaacag	aatttaaggt	aaatgacact	ggttaattta	aagtgcatac	4080
atgttgaaat	attgcagggc	tatcaagaga	agcacaaaag	gctctaagga	aatatacgtt	4140
gacaggtacc	atttcagtag	gcaagctaac	tgacaattta	accgctcacc	gaatgatagg	4200
tctcttaaac	attgctaatg	tagatgatgt	ttatgtttca	atctaatagg	gtctacggcg	4260
gtacagttgg	ctccagtacc	acaagtaatg	ggggatacag	atgctgctgg	agttagaaat	4320
ataatagaat	ctaacggata	ctttaaaggt	tcttctgaat	catggaactg	tgctgttgat	4380
ggtctaaatg	aatgtgatgt	tgcccttagt	gctcttggag	agctaattaa	tcatctgtct	4440
aggctaaagg	tgtgttggct	tgtttagttt	ttgcttttca	caaattaagc	aaaggaactt	4500
ttcataactt	acagtttcta	tctacttgca	gctagaagat	gtacttaagc	atggggatat	4560
ttttccatac	caagtttaca	ggggttgtct	cagaattgat	ggccagacga	tggtaaatct	4620
tgagatattt	aacaatagct	gtgatggtgg	tccttcaggc	aagtgcatat	ttcttttttg	4680
ataacttcaa	ctagagggca	gacatagaag	gaaaaattct	aatacttcgt	acggatctcc	4740
agtaagtaat	agccgatttt	tgtttaccta	tgtagggacc	ttgtacaaat	atcttgataa	4800
ctgtgttagt	ccaactggta	agcgactctt	aaggaattgg	atctgccatc	cactcaaaga	4860
tgtagaaagc	atcaataaac	ggcttgatgt	agttqaaqaa	ttcacggcaa	actcagaaag	4920

tatgeaaate actggecagt atetecaeaa aettecagae ttagaaagae tgeteggaeg	4980
catcaagtet agegttegat catcageete tgtgttgeet getettetgg ggaaaaaagt	5040
gctgaaacaa cgagtaagta tcaatcacaa gttttctgag taatgccttc catgagtagt	5100
ataggactaa aacattacgg gtctagctaa agactgttct ccttcttttg caatgtctgg	5160
ttattcatta catttctctt aacttattgc attgcaggtt aaagcatttg ggcaaattgt	5220
gaaagggttc agaagtggaa ttgatctgtt gttggctcta cagaaggaat caaatatgat	5280
gagtttgctt tataaactct gtaaacttcc tatattagta ggaaaaagcg ggctagagtt	5340
atttettet caattegaag cagecataga tagegaettt ecaaattate aggtgeeeat	5400
ctatctttca tactttacaa caaaatgtct gtcactactc aaagcaatgc atatggctta	5460
gateteaaet eacacceega ggateetaaa gggatttget tittatteet aatgttittg	5520
gatggtttga tttatttcta acttgaactt attaatcttg taccagaacc aagatgtgac	5580
agatgaaaac gctgaaactc tcacaatact tatcgaactt tttatcgaaa gagcaactca	5640
atggtetgag gteatteaca ceataagetg eetagatgte etgagatett ttgeaatege	5700
agcaagtete tetgetggaa gcatggeeag geetgttatt ttteeegaat eagaagetae	5760
agatcagaat cagaaaacaa aagggccaat acttaaaatc caaggactat ggcatccatt	5820
tgcagttgca gccgatggtc aattgcctgt tccgaatgat atactccttg gcgaggctag	5880
aagaagcagt ggcagcattc atcctcggtc attgttactg acgggaccaa acatgggcgg	5940
aaaatcaact cttcttcgtg caacatgtct ggccgttatc tttgcccaag tttgtatact	6000
cgttagataa ttactctatt ctttgcaatc agttcttcaa catgaataat aaattctgtt	6060
ttctgtctgc agcttggctg ctacgtgccg tgtgagtctt gcgaaatctc cctcgtggat	6120
actatettea caaggettgg egeatetgat agaateatga caggagagag taagttttgt	6180
teteaaaata ecaatteete gaactattta eteagatttt gtetgattgg acaaggtggt	6240
tttgcttttt tttaggtacc tttttggtag aatgcactga gacagcgtca gttcttcaga	6300
atgcaactca ggattcacta gtaatccttg acgaactggg cagaggaact agtactttcg	6360
atggatacgc cattgcatac tcggtaacct gctcttctcc ttcaacttat acttgttgat	6420
caacaaaaac atgcaattca ttttgctgaa acttattgat ttatatcagg tttttcgtca	6480
cctggtagag aaagttcaat gtcggatgct ctttgcaaca cattaccacc ctctcaccaa	6540

ggaacccgcg	tettacetae	gtgtcacctc	gaaacacatg	gcttgcgcat	tcaaatcaag	6600
acctgattat	caaccacgtg	gttgtgatca	agacctagtg	ttcttgtacc	gtttaaccga	6660
gggagcttgt	cctgagagct	acggacttca	agtggcactc	atggctggaa	taccaaacca	6720
agtggttgaa	acagcatcag	gtgctgctca	agccatgaag	agatcaattg	gggaaaactt	6780
caagtcaagt	gagctaagat	ctgagttctc	aagtctgcat	gaagactggc	tcaagtcatt	6840
ggtgggtatt	tctcgagtcg	cccacaacaa	tgcccccatt	ggcgaagatg	actacgacac	6900
tttgttttgc	ttatggcatg	agatcaaatc	ctcttactgt	gttcccaaat	aaatggctat	6960
gacataacac	tatctgaagc	tcgttaagtc	ttttgcttct	ctgatgttta	ttcctcttaa	7020
aaaatgctta	tatatcaaaa	aattgtttcc	tcgattataa	caagattata	tatgtatctg	7080
tcggtttagc	tatggtatat	aatatatgta	tgttcatgag	attggtcaag	agaaatactc	7140
acaaacagta	tattaagaag	gaaatatgtt	tatgcattaa	tttaagtttc	aagataaact	7200
gcaaataacc	tcgactaaag	ttgcaaagac	caaacacaaa	ttacaaaact	tataagactt	7260
aagttctgaa	ttccctaaaa	ccaaaaaaaa	aaacagaaca	tattttgttg	catctacaaa	7320
caacacaaac	ctacatagtt	tataacttac	tcatcactga	gattaacatc	agaatcattc	7380
tccatttctt	catetteact	ctcatcatca	tcaccaccac	catgatgatt	ctcctct	7440
tcacgtaacc	tagcaatctc	actctgagct	ctatcaacaa	tetgettett	ctgcaactcc	7500
aaatctctct	gaaaatcagc	tctcatcttc	tccaactcct	tcatttgctc	tttcttactc	7560
ttctccatct	tctcataaac	cttcccaaac	ctctcaacag	aatccgccaa	catcttatac	7620
gaagcagcgt	cattaacctt	cttcctctcg	tactcaacct	catcatcctc	atcctcctcc	7680
tcttcagaat	caccaggact	atccatcatc	tcatcaaacc	cattagactt	atctaaataa	7740
accttagtgt	tcataaacac	aaactcacct	gaatcaacac	cacaagctaa	acctaaatcc	7800
gacttgggcg	aaacacaaag	caacatatcc	aacttattga	aaaacgacca	tttactțgaa	7860
cctaaacctg	atttctcaac	cttaatcttc	tcttttctat	acttcctctt	caagtcatca	7920
atcattctcc	tacattgcgt	ctcagatttc	tccatcctta	gctcctcact	cactttctca	7980
gctacttcat	tccaatcctc	gttcctcaaa	ctccttctac	ccaattgcaa	aaacctatct	8040
cccaaactt	caagcaacac	aa				8062

Figure 1



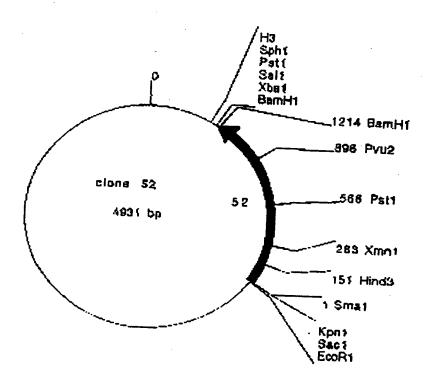


Figure 2

Commente/Asterences: 52= 3' side of S5 (AIMSH3) 1244bp in pUC18/Smat

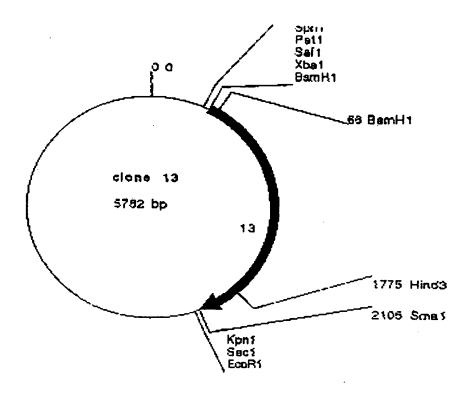


Figure 3

Commenta/References: 13 = 3' side of S5 (AtMSH3) 2104bp in pUC18/Smail

~	ij	CT AMGA	SPEC	2525	GAAA	WITG	S.	CCC.	AGTT(ದಿಂದ	atage	22.52	took!)KCCH(CAT	CATI	FFCT	CHIE	A.R.C.G	grargescscsabratisscarcerasticscertascicescaesercaescaestecatitetotibarcssassa	36
пг ж	GATTA		ucgaataaagcaatt	Jak 6C	artt.	A MIG	ပ္ပ ဗ္ဗာ	\$¥ \$	శై ర	X X	20	Ω Σ Φ	32K	att	TOT	F	FIC	TTC	, 15 et)) -	_ 독 2
145 16	A.P.A. K	00 a.	3 ×	700 %	9 0 0 0 0 0 0 0	BCT T	SAC SE	975 19	رد 9	744 2	553	61.8 8	500 4000 4000	क्षेत्र	ر بري دري	TCA S	ACA T	99 2 00		9)) a	9 50 8 50
205 36	A.M.G	ATA I	TCC S	ည် န	ACT	GTA V	700 8	TIC	r Ti	000 P	700 8	A A	CGT R	7. 7. S.	CII	CTC	TCC	6.4C	S S S S	CTC L	<u>લ</u> જ જું જ
265 565	Ç (3) ∢	8	900	A P	200	A A	A A	CC-1	X X	r T	TCT	CCT) () () ()	ACT T	3 00		(CC)		8 ~	GAT 0	23.24 25.24
325 76) 1	PAT	TTA L	ig ¥	n A	A CA	TTT	or T	S C	90.00 R	TYT	CTS C	5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5	သူ့	TCG			GAA E		GTT	38. 95
365 96	ပ္သ	684 8	₩ ₩	TCA S	S	S S	25G R	A.2.3 K	TAC Y	SCA.	గ్రో 🕹	T76 (सूत्र अ	950 C		57.5	CIG ×	9 4 6	CTA	A.A.G.	444 115
116	\$ S	AFG K	TAC X	50 *	GRT	5 5 5 5 5	GTT V	17G	8.7G	इंग्ल	. e.	Grr (100	TSC				TTC		4 5 5 5	504 335
505 136	মূত্র ক্ষু	GAC D	900 8	ស្ត្ ស្ត្	ATC	GCB A)) 4		67G	TTG (667 A	ATT 1	()	601 A	1. do	ATG C	GAT () (2) (3) (4) (4)		TIC	564 155
565 156	3 TG	₹C6	000 4	AGT	GTG V	3	5 r	FTT	CGR 1	17C /	ABT T	777.0 F	1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1	cts y	E CE	8GA C) 91.0	57.5 V		ភ្ជុំ ដូ	624 175
625 176	S G G	TEC	SAS X	ATT.		67.B	575		2000	r E	GP.A. 3	ACT C	ઇ જ ઇ જ	600 F	ATT A	AAG T X	TCC C	Car E		\$ C\$	684 195
865 196	24. 5	9 8) ₩ ₩	200		TTT	TTC (2000 R	GGA G	cre 1	rcg s	500 T	TTG T L		4 304 F	ASA G K	# 220 #			ਅਜ਼ ਤ	744 215
745 216	හ ජ	GCT.	9 SAC	GA:	ATA I	AGT (66T (567 1 6	16F G	0 000	GCT 5	្ត ស្ត្រ ម	5 4 5 3 5 6	COL T	TTT G	631 T		8 5 S	AGT A	AAT N	804 235
805 236	TTC	TTG	GIT 1	ToT C	GTT	GTG 4	GAT C	GAG A	AGA G R	GTT *	Augg Ti K	TCG G	GAG K	ACA T	ម្ ម្	F U				GA.A. E	999 255
965 256	₹ ¥	361 S	TIT (GAY	GTT 3	ACA R	្សា ខ្	0 691	GTT G V	GTT G V	66C G	្ត ក្រក្	K 489	FIT F	TCG A	ACA ACA G	GGT G	6A.A.G E v	617 G V V	GTT V	924 275

984 295	1044 315	1104 335	1164 355	1224 375	1284 395	3344 415	1404	1464 455	1524 475	1584 495	1644 515	1704 535	1764 555	1824 575
rtG L	ទាំវិទ	AGC S	TTP.	GTT V	CAT H	S S	AAT N	GTB	ATA I	TCC S	r T T E	ଜଣ୍ମ ଓ	ecr A	8 800 8
AGC T S L	17 G G	TTC A	AAC T N	aca g T	100 C	TTG T	7. A.	ACA G	TTC P	17.13 S S	GAG 7	ost g	Car G	ATG A
TTG A	FIF T	161 1	GGT &	1175 A	TTT T	rcf r	GTG A	CTT A	AAT T	AGT T	CCT G	5 5 7 7	ATC C	4 4 30 3
ATT T I L	rgg T X	GRT TO	ម ម្ជា ម	166 T	ACG T	CSC T	GTT G	RCR C	4 494 8	CAT A	2 2 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3	ATT C	GTT A V	rci s S E
GTG V	୍ର ଅଧିକ୍ର ଅଧିକ୍ର	ट्राइड	ក្ស ខ្	S	410 :	TTT	3 8 8	SEC.	CAT	s S	61'à V	GRT	ති වූ අ	245 a
5 5 5 5 7	¥ F-	TCP. S	्र सम	A TG	ည်း န	ر چې د	T;3 L	TAS.		150 150 19	att I	TCT	att I	చ్చ ే
9 29 19	5	in Cir	الان الا الان ال	S S S	ដូច	is 4.	9	8. 5.5	ج. در ال	≱7G N ·	602 A	TC?	11¢	K.S K
TIR L	1	(5) 4	म् स्ट्रिक	S A X	ပ (၁)	999	3 0 3	EAT M	CC 6	1GC	AG.22	क्ष ्ट्रम् प्र	(2) (2) (3)	ath I
GG. G	TC:	1 1 1	TGI	भूक उ	5 0	₹ 300	CJ 5	CAT H	r H	scr *	585 13	TCT S	న్ల .	၁၅၅ ၁၅၅ ၁၅၅
267 S	CT3	ត្ត ខ្	77% L	GCT *	GII	13C 7	ACT T	₹ F	TOTE T	101 S	TCT	atg K)) «	CTT L
AGE &	CC CC	1800 x	្ស ស	GCT A	14 KG	CTT.	rat X	다. 기.	CTG ×	kit 1	CC3 G	ecr s	¥.5 ×	CGG CTI R L 'igure
ATG M	CAG Q	GFT V	ATT I	ក្ ក្	ಭ್ರ	87C I	υ υ • •	100 3	9 (~ ¥	9 6 6	स्वरू अ	<u> </u>	GCT	CRG Q
11C 6	999	1) 6) z	GTT V	CTG	C 24	೨೨ ೫	٠٠ د د	ည ၁၉၅	Ç ≖	TCT S	445 14	113	ACT 7	rrt I
ZAT Z	נייד	ភ្នំ	ញ្ជុំ ភ	X X	i S a	सूर्य च	CTC •	₹¢₹ \$	AGA R	GTT	GPT	GTC V	\$ \$ \$ \$ \$	0 0
GAT	TTG L	T.	GAT D	a N	ATG M	TTT E	acr T	ig G	CII	SCT A	TTG	a S S	CAT	7.
AAT	CTC	CCJ	GTA V	ម្តីជ	AAC N	4 00 6	atg M	8 10 8	citi 1	C D D	S) E	TCC S	TTT F	8
TTC F	545 3	្ស ប៉ី ប	& \$05 ₹	3×	3 TG	TTT	CHC CHC	60 60 60	4 13 4 4	ĆIT L	agt S	orc L	ATC I	ეე ა
249	SCT A	A A	Z ×		ATT I	ე ე	₹		TCC	00 B	AGC S	GTG V	a R R	i i
45			GGT	GAT	ACA	ž Ž		35 8	gc _I	GCT	CTC	CIC	ACA T	TIP
TAT	3CA 5	4 4 5 4 7 8	AAT	445	CP.T	CIC	AGT	aat 2	TRI	TCT S	CAG CTC	TRT Y	ATA	ATT
	982										1585 496			

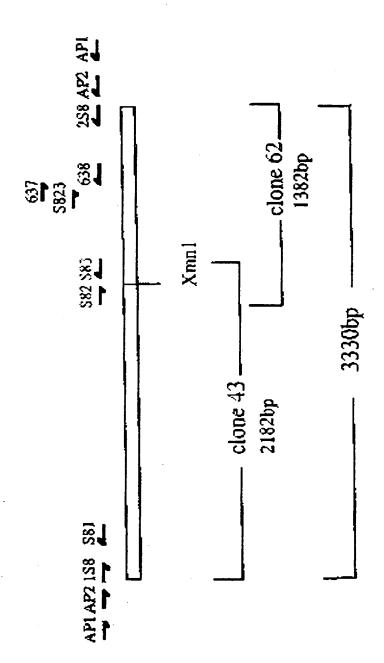
1884 595	1944 615	2004 635	2064 655	2124 675	2184 695	2244	2304 735	2364 755	2424 775	2484 ·	2544 815	2604 835	2664 855	2724 815
TCA S	A.	GCT A	TTE	TTC T	AFG K	ij G	AGT S	S R	ror	4 4 4 4	STC H	ATA I	ن ق	e 133
ATF	ე გ		77. S	CAT	30 F	CTA	11 11 11 11 11	TTC L	က္ခ ဖြစ်	TTA			CHA	
GTT	S S S S	S S S	GCT ▲	∯ ⊨		60T	AG₹ S	161 C	GAT	ATA I	450 450		CIT	144 C
rcr ss	XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX	1 1 1	ata I	ATC I		CJA	A AG	GAC	GTG V	act T	780 C		STC.	3
rt. I	AAT	TTT	5 2	ტ ცე	ers v	្ត ភូ ស	STC L	ក្ ក្រុក	TTT	ာ (၁) မ	TAT	60 K	CEC 6	AGT F
TYG	CTA	కే రీరా	T C C	TCG S	٠ دور ۲	CEST U	TIC	ල්ට අ	GRC E	oro oro	কু কুনু	GTT Y	CTG	a c
A. W.	\$ \$	ខ្លួក	GA7 0	: ::::::::::::::::::::::::::::::::::::	613 •	2	rgt S	139 ¥	رر ه	রুমু ১	ეცე ც	a So	A.A.G	G CAT GGC MG H G • R (Continued)
₹ ×	1C1) (1) (1)	្ស	S S	F %	 B 0	CAT	cr: L	cg7	5.4	i E	i) #)) (CAT (H (
TTS	CIC	TCC	K Service X	j,	1. 1. 2.	ទូឃ	¥	SCT A	GT¢ v	CAT	ಷ ಭ ಭ	ATC I	TTC	် ၁၅ ၁၅ ၁၅
CTT.	CTT	FC.F	स्टु इ	TTT	ATG H	د تا د تا	TCG S	ಕೈ ಚ	787 *	CGT R	CAT B	ક્ષ્યું ક્ર ક્ષ્યું ક્ર	TCB	ATC I
ACT.	χ. 	etc I	366 R	S. F. S.	E C	2. 1	GCT A	67 F	S. S.	SGT	113 L	TGC C	ಲ್ಲ ಕ	AGT S iguz
S	អ ទី ទ	CT.	ATC I	TTG	6) 7	S S		305 31	r S	TCT	ATT T	8 S	ర్ష్మ	្ល ប្រធ
200 K) (၁) (၁)	ata L	G7C v	A Taken) kg g	i i i	S Z	GC7	200	990 0	ACE T	AAG K	GT& V	S
ر د م	E EAT	O O O	Ē.	ණු ද	7 17 17 17)) () ()	ctc v	ASG K	त् 13 वं स	ata 1	0 P	G3A G	1.T.T F	50 K
ACT	G&C 0	្សា	GTT	ATT	GAT D	Q ¥	att 1	11 11 11	TCT	74. ²²	aas n	ð,	TCC	្ត ប្រជុ
ঠ ও ≰	GTT V	TTG	ec.	GCT ▶	617 >	Tr. Y	9 4	SAT	CTA 1	asa I	800 a	ATG S	i ci	A E
TCT S	GTG V	් රී ය	ž 0	ည်း) 20 20 20 20 20 20 20 20 20 20 20 20 20	<u>ئ</u> ئى ھ	rij 1	Ž.	ACT T	540 1	G3C V	200	GTT	8 %
20	GTT V	ត្ត ក្	250 8	AAG #	cje L	A1'T I	CAT B	75.C Y	TCA S	5 11	TTC	84	S C	T L
atc H	ដូ	S R	رن ه ون	A AG	in Se Se	RCT T	E SA	7× ×	ÇŢŢ	5	DE N	d GGP	GCT	TTC 7
AGT	T D	GTT V	த் நீ ப	ပ္ပိုင္က	ATA I	A.a.g K	j k	AGA R	700 5	2 G	GAT	ACC A	P.TG	CIT >
1825 576	1885 596	1945	2005	2065 656	2125 676	2185 696	2245 716	2305 736	2365 756	2425 776	2485 796	2545 816	2605	2665 B56

2725 876	S S	TTA L	AGT	GAA E	ಲ್ಲಿ ₹	S S	CAC B	ATA J	ATC 1	AGE A)) (161 J	TCT 18	FCT (CGT 1	TCG (CIT	CTT	AIA. I	rta L	2784 895
2785 896	GAT D	្សិស្តិ ខ្លួ	CTT L	a A A	A B B B B B B B B B B B B B B B B B B B	ည် ပ	PCT T	AGC S	ACA T	Ç ޱ	CAC O	0.00	GTA C	000 g	ATF (I	000 1	TRT Y	Į,	ACF.	IIA L	2844 915
2845 916	9 0 0	CAN	OHO L	CIB	425 8	GA.S	are X	ACH M	TGT	T 53	GTT (CTT 1	TTT (STC 3	200 c	CAT	TAC	CCT	3. A. B.	ATA	2904 935
2905 936	60T	945 E	A TC	AGT	a Ac	3 0	7.7C	100 PM	តិចារ	TCT	STT V	ა ეეც	ACE 1	TAC (C.S.3.	C CTC >	TCG S	TAT	CTG	ACA T	2964 955
29 65 956	TTG	9 9 0	X X	GAT O	* * *	သ္မ	S SGI	13.5 13.13	GAT	i E	GAT	GA: (GTG 3	32C ∓	7. 7. ×	CTR.	TRT	A A	CTT	GTG V	3024 975
3025 976	SGT T	GGT	CTT	S S	\$ s	A R	3 S	TTT	500	TTT	A A A G	511	SCI O	is cx	E CO)) () ()	වූ ය	ATA	CCF	100.th	3084 995
3085 996	708 S	TOT	rta I	ខ្លួ	ನ ಧಿ) (၂)	RT T L	ž v	₹ ¥ 2	50 %	j S	Sed X	TTG	45 m	5. CT	GAG	GT.A V	CGT A	5 4	AGA A	3144
3145 1016	SAG B	3.G3	A.A.T	aca T	Ω κ	at 6	ಕ್ಷ ೮ ೮	గ్ల (క్రి బ	ಕ್ಷ ಬ	ig ra f	ថ្ងីទ	CAT	Gist. (13 13 13 13 13 13	900	#G# #	ပ္ ပ္	4) (1)	ញ្ញ ស្វ	GAA E	3204 1035
3205 1036	TCT S	ATT I	s S	GCT	Ť.	GGT	G&C D	TTG L	17.L F	90 G	္ရွိမွ မ	CTG	. 2.9.7. K	Lil	TO A	ر ت ت	TCT S	E GAA	GAG E	G.P.C D	3264 1055
3265 1056	رن 1	TGG B	X A	ర్ట్ ∢	T'IC ₹	GAG E	ri f e	TT.A. L	reg K	C&7.	607 8	D N	25.25	ዶፓፕ 1	SCT.	299	ag ×	atc I	AGA R	cra i	3324 1075
3325 1076	ş Ş	ឡី 🎍	ACT	TGT	Į, s	fst E	35.	TTIE	ETCT.	Tarc	atta	TAGC	T) a a	4400	GGTC	TYGA	TCAT	CTGT	TRGT	TGA TITABICTIAACHITATAGGAGCTGCAAGGICTTGAICAICTGTIAGTIGCG	3397 1062
3398 1	TAC	TACTARCT'	F-4	ATG TC N	fct an	ATT AG	AGT HI S I	eta ec 1	ace ag T	AGA AEA R	ক্ষ ভ্রমত ন	े इ.स.	iat Tig Agag N	ر ا	λ 4 Σ	ATG &	GAT T	TCT A	AAS C N	ر د و 4	3458 5
ይ ይ ያ	GTG V	7 7G	5 to	TRC	ATC	TTT F	TCT S	ಸ್ವರ್	បួធ	603 4.	c cca ima ad a Fignre 4	IR AALESPAREARRAARAAAAAAAAA * ce 4 (Contioned)	spenspasekkus (Contioned)	ano.	Assas.	ተ <i>ዶ</i> ልጹ	ደ ዳዳ	Kara	Ę		3522 16

Figure 5

日	医心 工房 多物 征口 必要 商先 南京 阿罗 写在 有品 化土 自止 两份 途地 下水 死命 即时 時下 心气 伤气 皮马鱼目 丘头 男男 克斯 医甲 计中 目医 巴苏 节认	HER THE STANDARD BY THE STANDA
X X X X X X X X X X X X X X X X X X X	7 A	en oost tii terveen en en skateen seen seen seen seen ook en skateen ook en skateen seen ook en seen ook en ska Noober en skateen skateen seen ook een ook en skateen skateen ook ook ook ook ook ook en seen ook ook en skateen ook ook en skateen ook ook ook ook ook ook ook ook ook oo
MANS.AC.	4000	AT THE BEALT OF THE PARTY OF TH

Figure 6



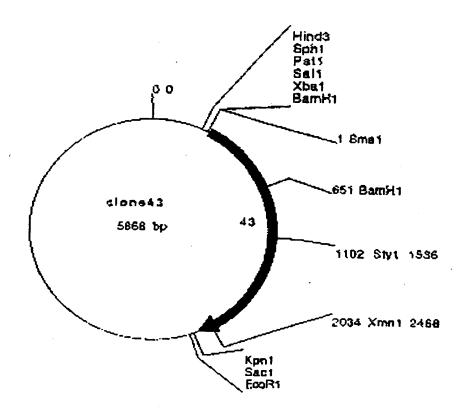


Figure 7

Communits/References: 43m 5' side of S8 (AIMSH6) 2182 bp in pUC18/Sm81

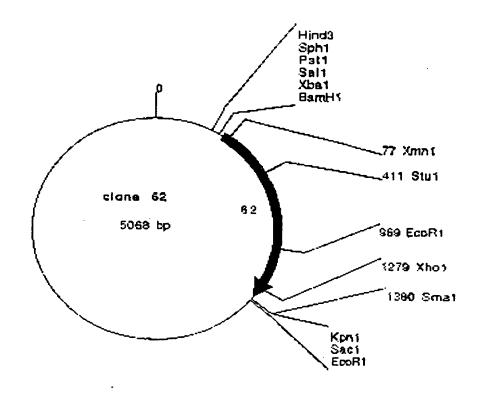


Figure 8

Comments/References: 62= 3' side of S8 (AtMSH6) 1379bp in pUC18/Smet

P. P	TIG)) (1)	CTG.	CCHG	XAT(STI	SCCC	CENT	(TCT.)	SCGAC	GCRA	ರಿದಿಲ್ಲಿ	S. P.	TTT	TGG	53	W3C1	ודומ	aaaagttgagooctgaggaatogtttoggocatttctacgraggggaaaatttttgggggraatctttccccc	80
TITGARITCICACACACALGATCTICICICACACTCTCACACAAAA AIG M	ទួ	본	TCAG	CTC.	MARK	CATC	37740	भटम	377.280	נוכונ	TCTX:	ACP.	rrcc	WW.	¥ Σ 5) # 940 8)) () ()	5 &	153
TCG ATT S I	14 14	Fe	776	TCT S	TTC	TTC	CRA O	ž Ž	25 ~	ACC 1	SCC A	55 et	#: [→	7 to	AAG K	র জ্	TTG	GTT V	TCC	213 24
GRT G	8 ≖	\$ 60.T	135 et	AGC S	ပ္ပ	9 9 9	ပြ ပြ	0 0 0	35.5C	50	69 69 69 69	9 1	766 3	TTT F	AAT N	GTG V	A.R.G	स् स्	800	273 44
GCT A	$\mathbf{z}_{\mathbf{z}}$	¥ ×	ပ္ပ ဖ	ည္သရ	€	TCT	5 4: 7:3	CG: R	TT. T.	CCT P CCT	GIT	JCG S	4 F	TCT \$	GTC •	GAT	6.2.G E	GTT	AGA	333
	() n	CAT	ij.	5 .	9 a.	9 6 9	AAG K	GTT	935 A	CGT R	cer R	7 63 C	رغو ت.	() () ()	rcr	ქ	TTT	æaG K	000 4	50 80 50 84 50 84
	<u> </u>	TCC	225	GST G	GAT D	GCT A	7C6	70C \$	crs L	TTC	3 S	rat N	art I	ž TČ	CAT #	AAG K	TTT	Gľa V	ž Ž	453 104
	רא פי	GAT	CGA R	GAT	TGI	TCT	3 50	្តិ ក	AGG R	ಸ್ವಿ 2	5 £	e G	GAT	513	GTT V	ပ္တည	71	aat B	GAT	513 124
		CTA	767	MATC	A X	00.T	म् । ।	GAT D	ر د ورد		122 ~	ရ ရ (၂)	<u>)</u> ,	R R	3 2 3	r g z	A LAT	, 6	AAA K	573 344
	(7) 1-7	A P	AGA R	R RC	C&T H	GCT A	TTT E	e; s>	ון ה	A SI	9 0	AGB R]] 2 4	ថ្មី ស	CIT	В.	TCA S	gta V	GAJ) E	633 164
ATA C		g g	GTA O	GAT	ည္သမွ	GAT D	67T 7	S S	00 th	600g	4 13 13	ACA T	5 0.	950	ATG 8	₩ 0.64	8 ~	CGT	GCT.	693 184
ეც 1		TTG	ARG (15 K	GTT	CIG	946 9	ÇKT O	ម មិន	ATG B	FCT.	TTT	X Ze.	GAG E	GAT	AAG K	GTP	r L	ស្ត្រ 	753
ည္မ		S	AAC I	H.F.	7.00 7.00 7.00 7.00 7.00 7.00 7.00 7.00	010 1	45. 75 F	atg N		C 25 C	58:	900	JIE	161 C	66A G	52.0	E A	A. A. A.	सूच्यु अ	813 224
		GEB E	5 5 5 5	ACC T	ች ት	TTT E	ដ ភូម	TGG N	CTT	. 549 3	S	TCT e	CG3	11.15	£66 ¥	GAT	8	RAT	r R	873 244
CGT C		ה ה ה	CAT (GAT	88.	CIT	TAC	CA1	A CA	15.50 X	ည္ညွ	TT	ပွဲ ဘို့ အ	ATA	CCA T	בכת י	GBT (GTT	TTC	933 264

Lyure

VSDOCID: <WO 9919492A2TI >

993 28¢	1053 304	1113 324	1173 344	1233	1293 384	1353	1413 424	1473 444	1533	. E831 484	1653	1713 524	1773 544	1833 564
C ATT	Tir get L s	GIT GGT V G	AAA GIT K V	ACT AIN I I	ATC GGG I G	TGT TCA	TCC ATC	rag gar K b	TAT ACG	aca gat T d	GRA TCA E S	GCA GAG G E	ATT TTT I F	CIT GAG L E
RTG GAC	त्रुप्त या स्र	្ត ទ	TAT Y	agt N	A. ×	KAG K	טפט טפט	CCA	A R R	GAT D	ICI	CTT	មិ ជ	AAT M
sar tat e Y	GAT GOG D M	TGC REAC	CST COM	SCT GCT	0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	CTA CAA L O	TCG CTT M V	GTT TCT V S	CTA AGG	atc ccc n c	GGT TCT G S	Act cct s a	Cet GGG H	atg gta N V
S. S.	GRG CTA E L	स्थ्य ४:55 अ	TIM GET L A	600 264 A R	GCA RGC A 5	0 kg E	T 7 T E	្ត ខ្លួ	[] [] []	GTA	A.A.	CTT L	₹ .	()
GTT AAG V K	म्बर र	© 7 G ∨	स्त्र प	3 4 ×	¥0. 	á z	TTG AGG	7.7	CAR 236	י ככא כאינ פייט אבט י	Tec Tit	orr ccc	GIA CII	MI GGC CAG AC G Q T (Continued)
rgg kst * s	Gag cro E i	agi get s e	Cas pag o k	೧೬೩ ೧೮.६ Q A	CCA TOP S S	Cag atc E i	GCT GCC	GCG ITA A L	Gragera Ear	CCA GTA P V	230 GG2	TGT CAT C D	GRA GAT E D	ش ⊂ ک
TAT X	TT TAT F Y	ACC ATG	GCA GTG	TCT GAC S D	TIR ACT	ate ara i x	TGT	CTT CGR	TCA AGA S K	TTG GCT 1. E	GAS TCT	ទី ទីធ	AAC CTA K L	CTC AGA ATT L R I Pigure
PAG CAS K 0	A.P.A. K	ATG B	3 3	F F	GTÀ ∨	£29 ₹	GTI	6CT	ť J	g o	eta I	C'ra bat L N	CTA L	TGT C
17.5 Q.v. S &	GTG GGG V	TGG AAG W K	ath gat I o	CTR GRA L E	GTT CRG V Q	ctr ctr e l	GCT TTT A F	TGT GCT C A	AAA 655 K G	GCG GTA A V	aat atg n	GLT GGT D G	TCT AGG S R	AGG GGT R G
ర్ట్ 🛪	rea X	GAC D	0 0	9 5 o	CT3 L	CAT H	TII	TCA S	AGT S	ACG T	AGA R	69.T	CTG	7 ¥ 7
ATG TCF	TIC TIT	GAG CTT	CAM MGT	ayc gag i e	AGG AAG R K	GOC GTC	TAT GGA X G	CAT GCA	181 GAC	GGG TCT G S	GCA GIT G V	TGF GCT	AAT CAT	CAL GIT
AAC AAA ATG K K M	Gre Cli v L	CAC AAG H K	atc ict i s	GGA CGA G R	ATT CCB	CCT GAT	ACT GTG T	AGC GAT	GTG TTA V L	TTG ACH	GCT GCT	TGG AAC	CTA ATI	CCA TAC P Y
									1474					

1893 584	1953 604	2013 624	2073 644	21.23 66.43	2193	2253 704	2313 724	2373	2133 764	2493 784	2553	2613 824	2673	2733
AAAC ¥	GAT O	r To s	S S K	s s s	GAT	AAA K	ರಿ ರಿ	ACT	CAC H	SCA G	ā.	155 5	att I	R R
GAT D	¥ ¥	हें इं	600 0	afta X	att I		8 4 8	ម្តី ធ	att 1	GCT	A.Y.A.	C&T D	AGC S	រុះ
CIT	CIC	ជ្ជ ភ	CTC	££3 K	4 5 5 5 5 5	CTC	र्द्ध व	GCT *	C CTC	TCT \$	ر تا در	38 🐔	ი მ რ	2 1
trt Y	P P	NAC SE	CTG 1	င် ၁	rgt S	3 5 %	77C F	Ž z	ម វង្គ ម	ည်သ	ZAT.	<u>უ</u>	AGT	ACT
₹ K	E E	50 %	A A	ر. اد	ર્જુ જ	TAT Y	3 0	ξ S	ICT S	₽Ç:] 5	C C	cit v	် (၁၅ %	5 5
TAC	i i	ACG J	4 15 w	C:7 L	۲۰: ۳:	ניני ג	TCT	0,41 □	70G	∱ ≈	SAT D	<u> </u>	AGA R	A × T
TTG L	a.fc I	بر 1	₹7.¥ 1.	() e:); };	ت ت	רים ני	\$ () €	နှ ပါ ၁	ects Section	¥03 ¥	Til E	AG# 8	. 430 6 9 9
#CC	7GC ₹		ည် သ		ξ ×	۶.: م	T : T F	STG Y	a FC:	1. 1.	GCT F	<u> </u>	90 de 1	ညီဗ မျှန်န
999	aat R		ያር የ	:: - -	:713 ¥	्ट हुन्	F			র্তু ব	CP.	S. F.	GAG E	ATG M G (C
S S	AGG R	GTT 4		SIG S	AT.	4 21 1: 1:	೧೩೪ ೧	S O	છું જ	TTT F	ភ្នំ ស	TGG B	ပ္ပ	A.B.C.
CCT P	T'ta L	41.9 *	ž×	S S	200	F K K K K K K K K K K K K K K K K K K K	ਹ ਹ	0 4 2	4 4 3 3	JCT S	ន៍ ខ្លួ	CT3 L	CTT L	CCA P Fig
ত্ত	CIC	5,57	CAC SS SS	000 K	ري دي دي	្ត្	505 13	ე ე		* 5 %	S 4	1. 1.00 1.00) 1	មិ ភូមិ
តិ តិ	CG#	g	arc r	40° 80° 80° 80° 80° 80° 80° 80° 80° 80° 8	TTT	e G	S S	7 (3:7	TTI	CTG L	TTT E	\$ 0	alta H	ACG T
627 0	* *	រូវ ប	Tat Y	Ę N	() ()	3 ×	ag x	i de z	ניז	STC V	ATT X	ATC I	GRI	7. 0.7.0
ក្នុក	ភូក្សា G	X X	200	₹ 8 æ	ag X	g g	ন মূ	CC.A	633 6	GAT D	GTT V	¥ ¥	ÆT ¥	178 1
AGC S	ACT T	A.A.T.	ပ္သမ	GTT	STT V	CT2 L	GT⊉ ✓	TYT	ATC.	CIP	CCT	CTT	9 0 0	TTG L
P.A.T N	ర్ట్లో	ATC	म्	နှင့် န	ජ ර අ	SCT A	ITA L	9 3	CTT L	ត្ត	AGG R	ATA	GIT	S S
AAC	act s	S SG	ATC	TCT	2 0 2 0	3 TTG	ata I	S S	A AIA I	နှင့်င န) ()	ور ال		9 9 4
TTT F	GTT V	S. E.	నే చి	X X	A A	TTG L	CCT P	GAT	g K	ATA L	atg K	925 G	1136 L	CCT P
ata I	ភ្ជិ	GTR V	ATG K	ATC 1	CJG L	CIG	CIT				AGC S			CAT
3834 .565	1894 585	1954	2014 625	2074 645	2134 665	2194 685	2254	2314 725	2374 745	2434 765	2494	\$524 808	2614 825	2674 845

2793 884	2353 904	2913 924	2973 944	3033 964	5 0 0 E	3153	3213 1024	3273 1044	3333 1064	3393 1084	3453	3521 5	3579 1.9	3606
1 1GC C	4 ACA	A ACT T	a Tac Y	C TTT	G ABA	2 CAC	a Gre v	೨೨५ ಕ	r ret	17 GCC A	55 TCT 8	CACTATCTGAAGCTCGTTAAGTCTTTTGCCTCTCT	AAA K	
101 8	ATG	8 8 8	3 50	CTC L	A CG	200	30	2 G	TCA S	TAR C	1 1CC S	ည၅	E 3	
SAG E	ATC	rat N	GRT	atg M	₽ ₽	GAT D	CIT	GCT	TIC	ASC N	AMA K	TIL	CGA	
167 C	AGA R	950	TIC	ეებ 8	6.7°C •	TGT	წ ა	GCT	943 .	T CAC	3 ATC I	AGTC	ည်း ည	
000 a	S o	נות ה	act T	ng C	CGT R	GGT G	TAC Y	con c	15.T S	339 4	9 9 9	GTT?	TTT (
GTG Y	rct	GIT V	AGT S	ಕ್ಟ ರ	80	CGT ★	၁၅ ဗ	30 S	400 800 800 800 800 800 800 800 800 800	oro v	53T	ופכנכ	TTG I	. ব
13.5°	ನ ನಿ	۲۰ در ۲۰	AC.T	STT V	ე გე ≍	رن 10 م	କ୍ଷ୍ଟର ଅ	स ् १०७	() 1	₹ 55 45 45 45 45 45 45 45 45 45 45 45 45	JCG K	TGa	1 4.3.7. 1 1	• Irme
160 C	ם מ	93 ×	430 5	ž×	1115 R	30	CCT	#C# 1	្ត ស្វ	TCT	TTR L	TATC	45.4 4.8 8.8	• (Continued)
ည ဗ	CIT	ಕ್ಷ ಕ್ಷು	AGP. R	ទូ <u>៥</u> ១ ឆ	00 ≪	T&T Y	16T	4 4 5).ეყ ა	arf I	, 2000	383	ATC A	. <u> </u>
CTT C	36G 38	្ត ខ្មា	္မွာ ဗ	GTA V	77G ₹	GAT D	GCT À	GT:	7C S	GGT G	TTT F	TAB	TAT A	igure 9
ã o	ACA Y	ACT T	CI3	CTG T	n Again	tcr s	SGA G	() ()	7.80 8.	GTT V	TT	€ 3. 1.	T ATT L	w
) () 4	TT E	160 C	6.58 8	CAC TAC	A. DA	2 653	១ ទី ធ	4.0 4.0	7.10	. 1]6	: #C3	. ATG	760 T	त्र भ च भ
7TT E	atc I	ក្នុ ស្	243 0	CGT R	P CC	ឯ្	J T) 12. 12.	A.R.C	E v	0 GAC	SCT A	7. 4. X	7 X
ATC 1	ACT T	GTA V	C.T.	7 1 L 6	CTC L	3 ×	TTR L	រូវ	ಕ್ಷ 200	2. X	1300 X	, ATG	# # * *	स्तुत्र <u>भ</u>
GTT V	155 0	7. 57. J	atc I	GII	oct P	TTC	CGT R	ata I	ر در در	្ត ក្	GAC O	TA.	CTT A	K X
g_{4}	ត្ត ភូមិ	TTT B	GTA.	₹ÇĞ S	ე ლ	₹ \$	TAC Y	ი გეე). - -	17.66 14	20	\$ ×	oct P	A.A.A.
CTG	ر د تا	T		TAC T	TAC	1 60	TTG	GCT A	ប្តី ។	SAC O	5.50 GBM 6.00 EBM	CCC M	ATT C	A A
TGT C	ည္သ	AGT	ర్ట్ ఇ	ក្ល ២ ៨	CAT H) (TTC	atg M	AGA A	ନୁ ମ ୟ			ب	r Ara
54		9 4 5 3	GRT D	ATT I	ACA T	atg K	इ. इ.	ည်သ	A DAG	C A	ATT	1GT C	RTG III M F	A X
	S. S. S. S.	4 15 15 15	9 5 a) () 4	4 4 4	25 ≠	CTA	800 4	ATG S	ក្ ភ	ည သ	TAC	G G	PAA K
734 865	794 885	\$ 24 208	2914 925	25 4.00 4.00 8.00	3034 965	3094 985	33.54 1005	3214 1025	3274 1045	3334	3394 1085	3954	3522	3580 20

Figure 10

	1 2 0	1
10 statement in the statement of the contraction of		A CR. 2 44 - 2 5 A 6 M 5 _ 3 2 6 2 K
1919 THE POST OF SAME SELECTION OF THE PROPERTY BY AND ALCOHOLOGY OF THE SECOND OF SAME SECOND OF SAME SAME SAME SAME SAME SAME SAME SAME	: G	A. 54.54.44.44.44.44.44.44.44.44.44.44.44.4
ada processing beginned by the contraction of the c	**	10 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
and appearance and the second contractions of the second contraction o	33	14 + A 7 9 7 9 8 9 4 4 4 9 9 9 9 9 9 9 9 9 9 9 9 9 9
est		14 - 34 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4
er der der der der der der der der der d		78-27-27-4
13 rutt der kyrder for der Endorder (f. Ede Chromes for de Chromes	# C + C + C + C + C + C + C + C + C + C	19.788224
11. 11. 11. 11. 11. 11. 11. 11. 11. 11.	13	14 - 14 4 4 5 A
21) Orse Markers and Fright of Correction of the	***	14. BB 6 3. A 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8
161 VERIOVOGRVERE EROME, ENTRE EROVERE EROPVERE EROPVERE EROPVERE EROPPERE		ACKERS-+1 RESC.YAAR
Hibrandellabathabathabathabathabathabathabathaba		**************************************
・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・・		Acaban-Fi Bord Lynaet

TTTTTTGGTTGCTAACAATAAAGGTATAGGGTTTTATGTCATCAATATAA	50
CTATATATAAAAGAARTGAAAGATATATTTTTTTTTTTTT	100
APPACARCAGAGACTTTTTTTTACTTTTTACATTEGTCARCAPAATACAA	150
CATAAACGACATCGTTTAATCATTTCCCAATTTTACCCCTAAGTTTAACA	200
CCTAGAACUTTCTCCATCTTCGCAAGCACAGCCTGATTAGGAACAGCTTT	250
ACCATTCTCATATTCCTGAACTACCTGAGTCCTCTCATTGATCTGTTTCG	300
UCAAATCCGUTTGTGACATCTTCTTCTCCAATCTCGCTTTCTGTATCATC	350
AACCTCACCTCTGCTTCACACGATCCATCGCCGCAGGCTCTGTTTCTTC	400
TTCCAGCTCTCTTCTTCACTCCCGGACCGCCGTAGATTCCCCCTTTT	450
TGTTCGAACCGGCATCGAATTTCTTAACCGTTTGAACCGCGACACCGTTT	500
CTCAGAGCTGCGTTAACCGCTTTCGGATCGCGTAGGTCTTGGCTCTTTTG	550
CTCAGAGCTGCGTTAACCGCTTTCGGCTCGCGCTCTTGTGTTACTGCTCCTG	600
GGTATCTGCTCGCCATCGTCGATGAATTGAGAAAAAGGAACAACGCGAAA	650
GGTATCTGCTCGBCATCGTCOATGAA.TGAGAAACGATGAAGATGAAGAATG	700
PTGTTGAGAGGATTETGATATTATATACGAAGATTGGTTTCTGGAGA	750
PTGTTGAGAGGATTETGATAL FARMATATACCTCTGGAACGTTCTTAGAGATG	30 0
ATTCCATCATCTTTTCCCA: 11100: 10100 ACC 10100 ACC 101010	850
ATTGACGACGTGTCATTATCTGATTTGCAGTTAACCAATGCTTTTTTGGGT TGGATTCGTGGTACACCATATTATCCGATTTGGGTCAATGGTTTTATATA	900
TGGATTCGTGGTACACCATATTATCCGA) TTGGCTCAATTAAATTAAACCA	950
AATTTGGTTTTCGGTTCGGTTATCACTTACACTTAAAATTTAAGCTAACCA	1000
AAAATTTTCGTAAAATTTATTTCGGTTTCAATTCGGATCCCTTACTTCCA	1050
GAACCGAATTATTCGAAACCGGGTTAGCGCAACCGAATACCAATGCCTG	1100
ATTEACTCGTTGGCTAGAAAGATCCAACGCTATACAATAATAGAACATAA	1150
ATCGGACGGTCATCAAAGCCTCAAAGAGTGAACAGTCAACAAAAAGTT	1200
GAGCCCTGAGGAGTATCGTTTCCCCCCATTTCTACGACGCCAAGGCGAAAAT	1250
TTTTGGCGCCAATCTTTCCCCCCTTTCGAATTCTCTCAGCTCAAACATC	1300
GTT1CTCTCACTCTCTCACAAATTCCAAAAATGCACCACTACTACGAAAATGCACCACTTTG	1350
CGATTTTGTCTTTCTAAAAACCCACGGCGCGACTACGAAGGGTTTG	1400
GTTTCCGGCGATCCTGCTAGCGGCGCGCCAGCGAGAGACCACCATTT	1450
AATGTGAAGGAAGGGCATGCTAAAGGCGACGCTTCTGTACGTTTTCCTGT	1500
TTCGAAATCTGTCGATGAGGTTAGAGGGAACGGATACTCCACCGAGAAGCGGT	1550
7FCCGCGTCGTCTGCCGTCTGGATTTAAGCCGGCTGAATCCGCCGGT	1400
GATGCTTCGTCCCTGTTCTCCAATATTATGCATAAGTTTGTAAAAGTCGA	282
TGATCGAGATTGTTCTGGAGAGAGGTACTAATCTTCGATTCTTAATTT	1700
TGTTATCTTTAGCTGGAAGAAGAAGAAGATTCGTGTAATTTGTTATTCGTT	1750
GEAGAGATTCTGATTACTGCATTGGATCGTTGTTTACAAATTTTCAGEAG	1600
CCGAGAAGATGTTCCCGCTGAATGATTCATCTCTATGTATG	1850
ATGATETTATTCCTCAATTTCGTTCCAATAATGETAAAACTCAAGAAAGA	1900
AACCATGCTTTTAGTTTCAGTGGGAGAGCTGAACTTAGATCAGTAGAAGA	1950
TATAGGAGTAGATGGCGATGTTCCTGGTCCAGAAACACCAGGGATGCGTC	2000
CACGTGCTTCTCGCTTGAAGCGAGTTCTGGAGGATGAAATGACTTTTAAG	2050
GAGGATAAGGTTCCTGTATTGGACTCTAACAAAAGGCTGAAAATGCTCCA	2100
GGATCCGGTTTGTGGAGAGAAGAAGAAGTARACGAAGGAACCAAATTTG	2150
ATTECTTCAGTCTCTCGAATCAGGGATGCCAATAGAAGACGTCCTGAT	2200
GATCCCCTTTACGATAGARAGACCTTACACATACCACCTEATGTTTTCAA	•

GARARTGTCTGCATCACARAGCAATATTGGACTGTTARGAGTGAATATA	2250
TGGACATTGTGCTTTTCTTTAAAGTGGTTAGTAACTATTAATCTAGTGTT	2300
CARTCCATTTCCTCAATGTGATTTGTTCACTTACATCTGTTTACGTTATG	2350
CTCTTCTCAGGGGAAATTTTATGAGCTGTATGAGCTAGATGCGGAATTAG	2400
GTCACAAGGAGCTTGACTCGAAGATGACCATGAGTCGTGTGGGAAAATGC	2450
AGACAGGTAAATTAGTTGAAACAACTGGCCTGCTTGAATTATTGTGTCTA	2500
TAARTTTTGACACCACCTTTTGTTTCAGGTTCGTATCTCTGAAAGTGGGA	2550
TAGATGAGGCAGTGCAAAAGCTATTAGCTCGTGGGTAAGGGAACCATCAT	2600
actitatggaattcetttactgctacttcggctaggatttaagaaatgga	2650
AATCACTTCAAGCATCATTAGTTAGGATCCTGAGAACTCAGGATGTTTTC	2700
TTATTEGTTATATAAYAAGTETTTTCATCAAGGAGTAACAAAAGTT	2750
BCACAATATTGTGTGTCACTGGCAAGGCATATATACCCAGCTAACCTT	2800
TGCTAGTTCACTGTAGTAACAGTTACGGATAATATATGTTTACTTGTATG	2950
TGGTACCCTCATTTTCTCTCTCATGGAGGCTTTGAAGCCTTGTGTTGAAA	2900
CTGGATAGTTACATATGCTTCCAACAGAACTAGCATGCAGATTCATATG	2950
CTTTCTTTTCTACTAATTATGTATTCACACACACTCGTTGTTTTTTTCAA	3000
AGATATAAAGTTGGACCAATCGAGCAGCTAGAAACATCTGACCAACCA	3050
AGCCAGAGGTGCTANTACTGTAAGTTTTCTTGGATAGGTGAAAGTC	3100
TTGCAGACTGTTTTCATCATTTCTTTTTTTTGTACATTACTTTCATGCTG	3150
TARTTAACTCAATUCCTATTCTGGTGTGATTATCAGATAATTCCAAGGAA	3200
GCTAGTTCRGGTATTRACTCCRTCARCAGCARGCGAGGGAAACATCGCGC	3250
CTGATGCCGTCCATUTTCTTGCTATAAAAGAGGTTTGTTATTTACTTATT	3300
TATOTTATCATGTTCAGTTCATCCAAGTCCTGAAAAATTACACTCTTCTT	3350
TACCARTCTTCCATCAAGCTGTGTAAAGGATTTGGAATTAGAAAATCATT	3400
ATTTGATCCTTTGTTTTATATCCAAGAGCTCCCTTGAAAAGATCTGTTT	3450
AAGATTETTTGCACTTGAAAAATTUAATCTTTTTAAGTGAATCCCCTACT	3500
TTCTTACAATGATCATAGTCTGCAATTGCATGTCAAGTAATATCATTCCT	3550
*STTACTSCATGCCCCTCTTTCTTAATGACCATTGTCTATGTTGTGTTTTG	3600
TETCGTETGCTGGAGAAATGATAGCTGATECAAGCTCTACATTA1'CATG	3650
ATTANGTAGCTGCTCAGGAATTGCCTTTGGTTACATTGCCTAATGGTTTG	3700
ATGTCAATTTTCTTCTGAATCTTTATTTTAGATCAAARTGGAGCTACAA	3750
ANGTOTTCANCTOTGTATGGATTTGCTTTTGTTGACTGTCCTGCCTTGAG	008E
GTTTTGGGTTGGCTCCATCAGCGATGAIGCATCATGTGCIGCTCTTGCAG	3850
CGTTATTGATGCAGGTAAGCAAGTGTATTCTGTATCTTATGTGTACCATG	3900
TGACTTCCTGTGCATATATTTGGGTTGCAGGAACTAATTCTGAATCACCA	3950
TTTGGTATGTTTTTTCCAGGTTTCTCCAAAGGAAGTGTTATATGACAGTA	1000
AAGGTAAACTGCTTGTATCGCCAGTTGTTTTGTTAAACAGAATTTAAGGT	4050
AAATGACACTGGTTAATTTAAAGTGCATACATGTTGAAATATTGCAGGGC	4100
TATCANGAGAAGCACANAAGGCTETAAGGAAATATACGTTGACAGGTACC	1150
ATTTCAGTAGGCAAGCTAACTGACAATTTAACCGCCCACCGAATGATAGG	4200
TETETTAAACATTGCTAATGTAGATGATGTTTATGTTTCAATCTAATAGG	1250
GTCTACGGCGGTACAGTTGGCTCCAGTACCACAAGTAATGGGGGATACAG	4 300
ATGCTGCTGGAGTTAGAAATATAATAGAATCTAACGGATACTTTAAAGGT	4350
TCTTCTGAATCATGG&ACTGTGCTGTTGATGGTCTAAATGAATGTGATGT	1400

* -

. 2

	4450
1GCCCTTAGTGC1CT1GGAGAGCTAATTAATCATCTGTC1AGGCTAAAGG	4500
TOTOTTGGCTTGTTTAGTTTTTGCTTTTCACAAATTAAGCAAAGGAACTT	4550
TTCATAACTTACAGTTTCTATCTACTTGCAGCTAGAAGATGTACTTAAGC	4600
ATGGGGATATITTCCAIACCAAGITTACAGGGGTTGICICAGAAITGAT	4650
CCCCAGACGATGGTAARTCTTGAGATATTTAACAATAGCTGTGATGGTGG	
TCCTTCAGGCAAGTGCATATTTCTTTTTTGATAACTTCAACTAGAGGGCCA	4700
GACXTAGAAGGAAAATTCTAATACTTCGTACGGATCTCCAGTAAGTA	4750
AGCCGATTTTCTTTACCTATGTAGGGACCTTCTAAAATATCTTGATAA	4900
C1G1GTTAG1CCAAC1GG1AAGCGAD1C1TAAGGAA11GGATC1GCCATC	4850
Cactcaragatgtagaaagcatcaataaacggcttgatgtagttgaagaa	4900
TTCACGGCAAACTCAGAAAGTATGCAAATCACTGGCCAGTATCTCCACAA	4950
acticcagactiagaaagactectccgacgcaycaagictagccticgat	5000
Catcagcctctgtgttgcctgctcttctggggaaaaagtgctgaaacaa	5050
Coagtaagtatcaatcacaagttttctgagtaatgccttccatgagtagt	5100
ATAGGACTAAAACATTACCGGTCTACCTAAAGACTCTTCTCCTTCTTTTG	5150
CAAFGTCTGGTTATTCATTACATTTCTCTTAACTTATTGCATTGCAGGTT	5200
AAAGCATTTGGGCAAATTGTGAAAGGGTTCAGAAGTGGAATTGATCTGTT	5250
GTTGGCTCTACACAAGGAATCAAATATGATTGAGTTTGCTTTATAAACTCT	5300
GTAAACTTCCTATATTAGGAAAAAGCCCCCCATGATTATTTTTTTT	, 5350
CARTTEGAAGCAGCCATAGATAGCGACTTTCCAAATTATCAGGTCCCCAT	5400
CTATCTTCATACTTTACAACAAAATGTCTCTCACTACTAAAACAAAACAAATCC	. .5450
ATATGGCTTAGATCTCAACTCACACCCCCAGGATCCTAAAAGGGATTTGCT	5500
TTTTATTCCTAATGTTTTTGGATGGTTTGATTTATTTCTAACTTGAACTT	5550
ATTAATCTTGTACCAGAACCAAGATGTGACAGATGAAAAACGCTGAAACTC	5600
TCACAATACTTATCGAACTTTTTATCGAAAGAGCAACTCAATGGTCTGAG	5650
GTCATTCRCACCATAAGCTGCCTAGATGTCCTGAGATCTTTTGCARTCCC	5700
AGCARGTCTCTCTGCTGCAAGCATGGCCAGGCCTGTTATTTTTCCCGAAT	5750
CAGAACCTACAGATCAGAATCAGAAAACAAAAGGGCCAATACTTAAAATC	5800
CAAGGACTATEGCATCCATTTGCAGTTBCAGCCGATGGTCAATTGCCTGT	5950
?CCGRATGATATACTCCTT6GCGAGGCTAGAAGAAGCAGTGGEAGCATTC	5900
ATCCTCGGTCATTGTTACTGACEGGACCAAACATGGGCGGAAAATCAACT	5950
CTTCTTCGTGCAACATGTCTGGCCGTTATCTTTGCCCAAGTTTGTATACT	6000
DGTTAGATRATTACTCTRTTCTTTGCAATCAGTTCTTCRACATGAATAAT	6050
PAATTCTGTTTTCTCTCTGCAGCTTGGCTGCCTACGTGCCGTGTGAGTCTT	6100
GCGAAAYCTCCCTCGTGGA1ACTATCTTCACAAGGC11GGCGCA5C1GAT	6150
AGAATCATGACAGGAGAGGTAAGTTTTGTTCTCAAAATACCAATTCCTC	6200
GRACTATTTACTCAGATTTGTCTGATTGGACAAGGTGGTTTTGCTTTTT	6250
TTTAGGTACCTTTTTGGTAGAATGCACTGAGACAGCGTCAGTTCTTCAGA	6300
ATGCAACTCAGGATTCACTAGTAATCCTTGACGAACTGGGCAGAGGAACT	6350
AGTACTTTCGATGGATACGECATTGCATACTCGGTAACCTGCTCTTCTCC	6400
TTCAACTTATACTTGTTGATCAACAAAAACATGCAATTCATTTTGCTGAA	6450
ACTTATTGATTTATCAGG?TTTTCGTCACCTGGTAGAGAAAGTTCAAT	5500
GTCGGATGCTCTTTGCAACACATTACCACCCTCTCACCAAGGAATTCGCG	6550
TCTCACCCACETGTCACCTCGAAACACATGGCTTGCGCATTCAAATCAAG	6600
Figure 11 (Continued)	And the second

atctgattatcaaccacgtggttgtgatcaagacctagtgttcttgtacc	6650
GTTTAACCGAGGGAGCTTGTCCTGAGAGCTACGGACTTCAAGTGGCACTC	670D
atggctggaataccaarccaagtggttgaaacagcatcaggtgctgctca	6750
agccatgargagatcaattggggaaaacttcaagtcaagt	6800
CTGAGTTCTCAAGTCTGCATGAAGACTGGCTCAAGTCATTGGTGGGTATT	6850
TCTCGAGTEGCCCACAACAATGECCCCATTGGCGAAGATGACTACGACAC	6900
*TTGTTTTGCTTATGGCATGAGATCAAATCCTCTTACTGTGTTCCCAAAT	695D
AAATGGCTATGACATAACACTATCTGAAGCTCGTTAAGTCTTTTGCTTCT	7000
CTGATGTTTATTCCTCTTAAAAAATGCTTATATATCAAAAAATTGTTTCC	7050
TCGATTATAACAAGATTATATGTATCTGTCGGTTZAGCTATGGTATAT	7100
AATATATGTATGTTCATGAGATTGGTCAAGAGAAATACTCACAAACAGTA	715D
TATTAAGAAGGAAATATGTTTATGCATTAATTTAAGTTTCAAGATAAACT	7200
GCAAATRACCTCGACTAAAGTTGCAAAGACCAAACACAAATTACAAAACT	7250
TATAAGACTTAAGTTCTGAATTCCCTAAAACCAAAAAAAA	7300
TATTTTGTTGCATCTACAAACAAAACCTACATAGTTTATAACTTAC	7350
TCATCACTGAGAT TAACATCAGAATCAT TCTCCATTTCTTCATCTTCACT	7400
CTCATCATCACCACCACCATGATGATTCTCCTCCTCTCACGTAACC	7450
TAGCAATCTCACTCTGAGCTCTATCAACAATCTGCTTCTTCTGCRACTCC	7500
AARTCTCTCTGAAAATCAGCTCTCATCTTCTCCAACTCCTTCATTTGCTC	7550
TTTCTTACTCTCCATCTTCTCATAAACCTTCCCAAACCTCTCAACAG	7600
AATCCGCCAACATCTTATACGAAGCAGCGTCATTAACCTTCTTCCTCTCG	7650
1ACTCAACCTCATCATCCTCATCCTCCTCTTCAGAATCACCAGGACT	7700
ATCCATCATCATCAAACCCRTTAGACTTATCTAAATAAACCTTAGTGT	7750
TCATAARCACAAACTCACCTGAATCAACACCACAAGCTAAACCTAAATCC	7800
GACTTGGGCGAAACACAAAGCAACATATCCAACTTATTGAAAAACGACCA	7850
TTTACTTGAACCTAAACCTGATTTCTCAACCTTAATCTTCTCTTTTCTAT	7900
ACTTCCTCTTCAAGTCATCATCATTCTCCTACATTGCGTCTCAGATTTC	7950
TECATECTTAGET ECTEACTCACTTTETCAGETACTTCATTCCAATCCTC	0000
GTTCCTCARACTCCTTCTACCCAATTGCRAAAACCTATCTCCCCAAACTT	9050
CAAGCAACACAA	BD62
	-

Figure 11 (Continued)

DOCID: <WO__ 9919492A2TI >

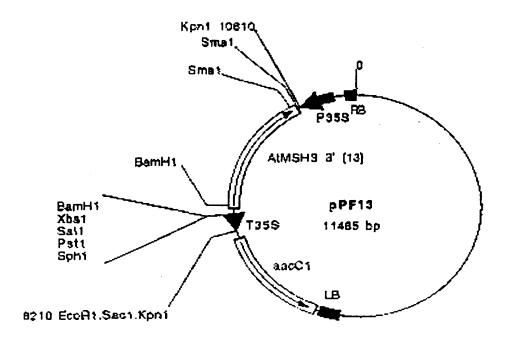


Figure 12

Comments/References: AtMSH3 3' side antisense : AtMSH3 3' (19 = 2104bp) from pUC18/13 Sal1/Sat1/T4 Into pCW184 BernH1/T4 In Agrobacterium LBA44O4

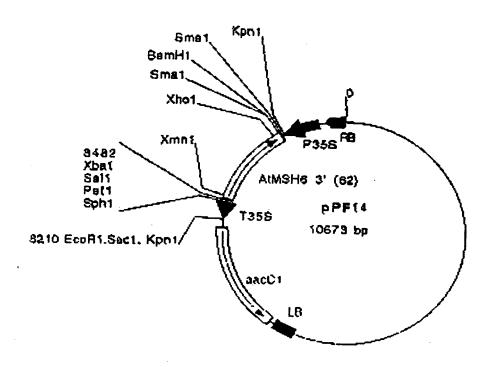


Figure 13

Comments/References: AtMSH6 (S8) 3' side antisens : 62 Sal1/Sat1/T4 (1379bp) into pCW164 BernH1/T4

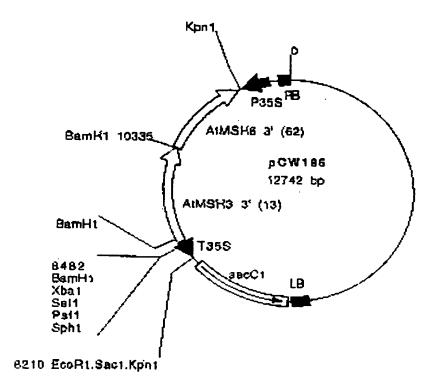


Figure 14

Commonts/References: AtMSH6 3'/AtMSH3 3' antisense : AtMSH6 (S8) 3' side (62=1379bp) Salt/Satt/T4 Into pPF13 (pCW184 AtMSH3 (S5) 3' side (13=2104) antisens)/Smat. in LBA44 $\overline{04}$

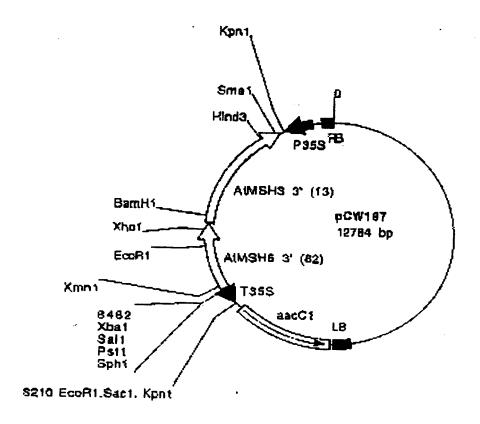


Figure 15

Comments/References: AtMSH3 3/AtMSH6 3' antisans (0): AtMSH3 (35) 3' side (19=2104bp) Sal1/SstVT4 into pPF14 (AtMSH8 (35) 3'side (62=1379bp) antisanse into pCW184)/Sma1. In LBA4404

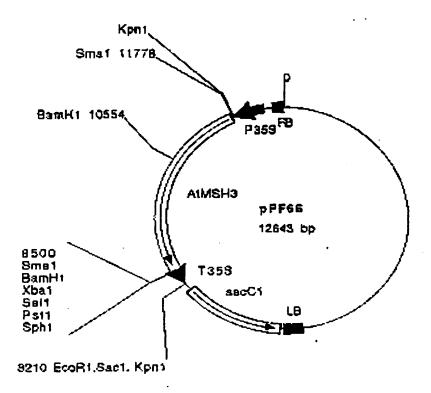


Figure 16

Commente/References: AtMSH3 (SB) complete, sease orientation : pPF28 (3342bp) Small into pCW164 Small

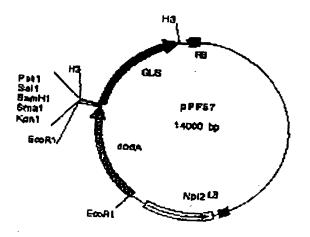


Figure 17

Commenta/Reforences: pPZP111 with code ExoRt cases in EcoRt site and Hinda GUS cases in Hinda site. Kenfl. All games under Promotor/Rempirator 353

Figure 18

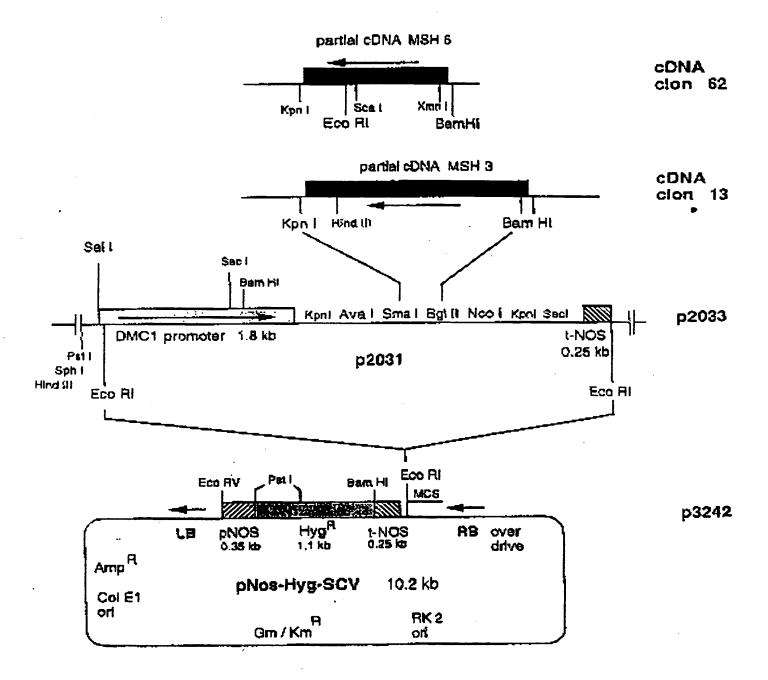
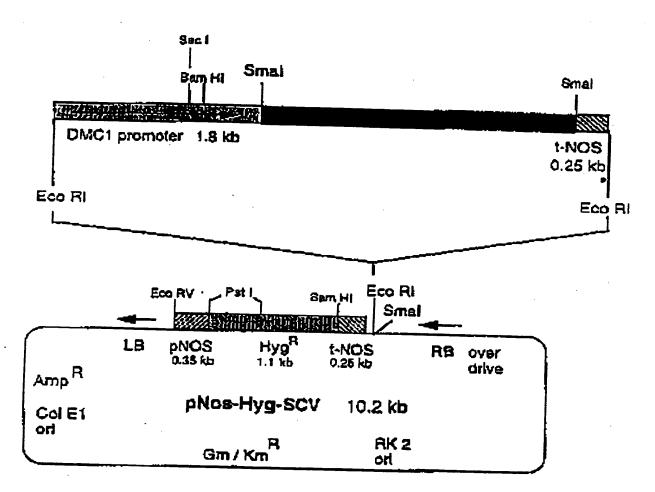


Figure 19

p3243



```
SEQUENCE LISTING
 <110>
               Rhone-Poulenc Agro; Betzner, Andreas Stefan; Doutriaux,
               Marie-Pascale; Freyssinet, Georges; Perez, Pascual.
               Methods for obtaining plant varieties
 <120>
 <130>
               395498C
 <150>
               PO9745
 c151>
               1997-10-10
 <160>
               93
 <210>
               Ł
               33
 <211>
 <212>
               DIVA
 <213>
               Artificial sequence
<220>
 <221>
               modified_base
 c222>
               11
 <223>
 <220>
 <221>
               modified_base
 <222>
               14
 <223>
               1
 <320>
 <221>
               modified base
 <222>
               17
 <223>
               I
 <220>
 <223>
               Degenerate oligonucleotides UPMU used to isolate AtMSH3 and
               ALMSH6.
 c300>
 <301>
               Reenan and Kolodner
 <302>
               Genetica
 <303>
               132
 <306>
               963-973
 <307>
               1992
 <400>
               1
```

etggatecae nggneenaay atg

23

3.5

<210> 2 <211> 23 <212> DNA

VSDOCID: <WO___9919492AZTI_>

```
<213≥
              Artificial sequence
<220×
<221>
              modified base
<222>
              15
<223>
              I
<220>
<221>
              modified base.
<2222
              19
<223×
              I
<220>
             Degenerate oligonucleotides DOMU used to isolate AtMSH3 and
<223>
              Atmsh6.
<300>
<301>
              Reenan and Kolodner
c3025
             Generics
<303>
              132
<306>
              963-973
<307×
              1992
<400>
              z
stggateert artgngthre raa
                                                                        23
<310>
              3
<211>
             24
<212>
             DNA
<213>
             Artificial sequence
:320×
<223>
             MSR3 specific primer 636 for PCR using cDNA of Arabidopsis
             thaliana ecotype Columbia
<400>
             3
tgctagtgcc tcttgcaage tcat
                                                                        24
<210>
             4
<211>
             27
<212>
             DNA
<213>
             Artificial sequence
<220>
<223>
             Primer AP1 for PCR using cDNA of Arabidopsis thaliana ecotype
             Columbia containing adapter sequences ligated to both its
             enda
<400>
```

ccatcctast	acqueteact atagggc 2	7
	_	
<210>	5	
<211>	33	
<212>	DNA	
<213>	Artificial sequence	
<220>		
<223>	Primer AP2 for FCR using cDNA of Arabidopsis thalians eco	Evpe
	Columbia containing adapter sequences ligated to both its	
	ends	
<400>	5	
actractata	gggetegage ggc 2	3
<210>	6	
<211>	30	
<212>	DNA	•
<2135	Artificial sequence	
	ar one a second	
<320>		
<2235	MSH3 specific primer S525 for PCR using cDNA of Arabidops	is
	thaliana ecotype Columbia	
	·	
<400>	б	
	HARMENTON	o.
aggittetgat	thigigigae gottsactta 3	·
<210>	Ť	
<211>	39	
<212>	DNA	
<213>	Artificial sequence	
<220>		
<323>	MEH3 specific primer S51 for PCR using cDNA of Arabidopsi	.9
	thaliana ecotype Columbia	
	_	
<400>	7	
	orewelltin agternage	9
agerragaca	ctgggttttg agtgtgagg	
<210>	8	
<211>	24	
<212>	DNA	
<213>	Artificial sequence	
<220>		
<223>	MSH3 specific primer 635 for PCR using cDNA of Arabidopsi	is
	thaliana ecotype Columbia	

```
<400×
                3
  gracgigett gaiggigtic toac
                                                                          24
  <210>
  <211>
               28
  <212>
               DNA
  <213>
               Artificíal sequence
  <220>
  <223>
               MSH3 specific primer S523 for PCR using cDNA of Arabidopsis
               thaliana ecotype Columbia
  <400>
               9
  tragaragta trragrateg ragaagta
                                                                         28
 <210s
               10
 <211>
               33
 <212>
               DNA
 <213>
              Artificial sequence
 <220>
 <223>
              MSH3 specific primer 1S5 for PCR using cDNA of Arabidopsis
              thaliana ecotype Columbia
 <400>
              10
 atecegggat gggesageas aageageaga ega
                                                                         33
 <210>
              11
 <2115
              27
 <212>
              DNA
<213>
              Artificial sequence
<220>
<223>
             MSH3 specific primer S53 for PCR using cDNA of Arabidopsis
             thaliana ecotype Columbia
<400>
             11
gacaaagago gaaatgaggo coottgg
                                                                        27
<210>
             12
<211>
             1250
<212>
             DNA
             Arabidopsis thaliana acotype Columbia
<213>
<2235
             Clone 52
```

<400> 12

cccdddwcda	дсардсавав	gcagcagacg	acttotegtt	tettegetee	cdaacccasa	60
terregarte	асдавссдав	teeggtagee	gaateateaa	caccyccacc	gaagetatcc	120
gccactgtat	ecttotetee	ttccaagegt	aagottotot	ccgaccacet	càccaccacà	180
toacccaasa	agcotaaact	ttotootoae	DCCCAARACC	cagtacccga	tcccaattta	240
caccaaagat	ctcccagag	atttctggsa	ccctcqccgg	aggaatatgt	tecegaaacg	300
ccatcatcga	ggaaatacac	accattggaa	cagcaagtgg	tggagctaaa	gagcaagtac	360
ccagatgtgg	ttttgatggt	ggaagttggt	tacaggtaca	gattettegg	agaagacgcg	420
gagatogcag	cacgegtgtt	gggtatttac	geteatatgg	atcacsattt	cstgscggcg	480
agtgtgccaa	cattlegatt	gaausuddat	gtgagaagac	tggtgaatgc	aggatacaag	540
atcggtgcag	cgaagcagac	tgaaactgca	gccartaagt	cacatggtge	алассдовес	600
ggccettttt	tcoggggact	gtcygnylly	tataccasag	ccacgcttga	ageggetgag	660
gatataagtg	gragergegg	tggtgaagaa	ggttttggtt	cacagagtaa	tttettggtt	720
tgtgttgtgg	atgagagagt	taagtoggag	acattaggct	gtggtattga	aatgagtttt	730
gatgttagag	toggtgttgt	tggcgttgaa	atttcgacag	gtgaagetgt	tcatgaageg	840
ttcaatgata	actteatgag	aagtggatta	gaggetgtga	ttttgagett	gtcaccagct	900
gagctgttgc	ctggocagcc	totttcacaa	casact gaga	agttitttggt	gycacatyct	960
ggacctacct	caaacgttcg	agtggaacgt	gcctcactgg	attgtttcag	caacggtaat	1020
gcagtagatg	aggttatttc	attatgtgaa	aasatcagcg	caggtaactt	agaagatgat	1080
aaagaaatga	ağttådəâdc	tgctgaaaa	ggzatgtett	gcttgacagt	tcatacaatt	1140
atgaacatge	cacatrigac	tgttcaagcc	ctegecetaa	cgttttgcca	tctcaaacag	1200
tttggatttg	aəaggateet	ttaccaaggg	gestsættte	getetttgte		1250

<210> 13
<211> 34
<212> DNA
<213> Artificial sequence
<220>

42235 MSH3 specific primer 255 for PCR using cDNA of Arabidopsis thaliana ecotype Columbia

<4002	13			
atecoggyt	c aaaatgaaca agtoggoott agto			34
_				
₹210∍	14			
<211>	27			
<212>	ÇNA			
<213>	Artificial sequence			
<220>				
<223>	MSH3 specific primer ssz for o	CB water was		
	MSH3 specific primer \$52 for p thaliana ecotype Columbia	CK ESING CON	A OI Arabidop	eis.
<4DO>	14			
gecacatete	actgiteasg coetege			27
<210>	15			
<211>	2110			•
<212>	DNA			
<213>	Arabidopsis thallana ecotype C	all combain		
د2235	Clone 13	erumpia		
<400>	1.5			
gccacatetg	actotteaso dectegreet saegttitge	catotoaac a	igtttggått	50
tgaaaggatc	ctttaccaag gggcctcatt tcgctctttg	tcaagtiaca (agagatgac	120
teteteaged	anthetetge ascagttgga ggttgtgkaa	aalaatteag s	itggatcgga	180
atctggetce	ttattecata atatgaatca caeaettaca	grataeggtt c	eaggettet	240
tagacactgg	gugacucate cuctaugega tagacatung	statetgete q	gcttgatgc	300
tgtttctgag	attretgett geatgggate teatagtret	toccagetca g	cagtgagtt	360
ggttgaagaa	ggttctgåga gagcaattgt atcacctgag	tittatotog t	gatatacta	420
agtettgaca	getatgteta gateatetga tatteaacgt	ggaataacaa g	aatctttca	480
toggactget	asagccacag agttcattgc agttatggaa	gctattttac t	rgcggggaa	540
gcaaattrag	cggcttggca taaagcaaga ctctgaaatg	aggagtatgd a	atctgcaac	600
	actettttga gaaaattgat ttetgttatt			660
caatgccgga	aaacttetet etgeeetaaa taaggaageg	getgttegag g	tgacttgot	720
cgacatacta	atcacttcca gegaceaatt teetgagett	gctgaagctc g	ccaageagt	780
tttägtcatc	agggaaaagc tggattcctc gatagcttca	tttcgcaaga a	gcccgctat	840

togaaatttg	gaatttcttc	aagtgtcggg	gatcacacat	ttgatagagc	tgcccgttga	900
trocaaggto	cctatgaatt	gggtgaaagt	aaatagcacc	aagaagacta	thogatatoa	960
cocccagaa	acageagetg	gcttggatga	gctagcceta	gcaactgaac	arcttgccat	1020
tgtgaaccga	åcttcätäåä	atagtttcct	caagagtttc	agcagatact	acacagattt	1080
taaggotgoo	gtteaagete	ttgctgcact	ggactgtttg	cactcccttt	caactctatc	1140
tagaaacaag	aactatgtcc	gtecegagtt	tgtggatgac	tgtgaaccag	ttgagataaa	1200
catacagtet	ggtrgtcate	ctgtactgga	gactatatta	caagataact	tegteccaaa	1260
tgacachatt	regeatgeag	aaggggaata	ttgccaaatt	atcaccggae	ctaacatggg	1320
aggasagagc	tgctatatcc	gtcaagttgc	tttaatttcc	ataatggctc	aggttggttc	1380
otttgtacca	gcgtcattcg	ccsagetgea	cgtgcttgat	ggtgttttca	ctcggatggg	1440
tgetteagae	agtatccagc	atggcagaag	tacettteta	gaagaattaa	gtgaagegte	1,500
agacateatc	agaacntgit	cttctcgtts	gettgtcata	ttagacgagc	trggaagagg),560
cactagcáca	<u> </u>	sagocasugo	ctatgoaaca	ttacagoato	toctageaga	1620
aaagagatgt	ttggttettt	ttgtczcgca	ttaccctgaa	atagetgaga	tcagtaacgg	1630
atteccaggt	cctgttggga	cataccatgt	ctcgtatctg	acattgcaga	aggataaagg	1740
cagttatgat	catgatgatg	cgacctacct	alataageet	gtgcgtggtc	tttgcagcag	1900
gagetttggt	tttaaggttg	oteagettge	ccagatacct	ccatcatgta	tacgtcgage	1860
catttcaatg	gctgcaaaat	tggzagetga	ggcacgtgca	agagagagaa	atacacgcat	1920
ддадаасса	gaaggacatg	aagaaccgag	aggegeagaa	gaatctattt	cggctctagg	1980
tgacttgttt	gcagacctga	aatttgctct	ctctgaagag	gacecttgga	aagcattcga	2040
gttcttaaag	catgettgga	agattgetgg	caaaatçaga	ctaaaaçcaa	cttgttestt	2100
ttgacecggg						2110

<210>	16
<211>	29
<212>	DNA
<213>	Artificial sequence
<220>	
<223>	MSH3 specific primer S51 for PCR using cDNA of Arabidopsis thaliana ecotype Columbia
<400>	16

9	gatc	à d&r	a ct	ggt	tttg	agt	gtga	33				*			,		29
<:	210,		;	17													
٤.	?11>		3	9 0													
<:	12>		r	NA													
<2	213>		X	Artif	icia	l se	quer	ice					•				
< 2	20>																
< 2	23>		M	SH3	apec	ific	pri	mer	3525	for	- PCF	≀ цэј	ing (ONA	ot a	rabid	opais
< 4	00>			7 7	апа	ecoc	уре	COlu	mbia	Į.							~~~~
зg	gtto	tgat	tat	gtgt	gac	gctt	Cact	ta									30
<2	10>		1	8							•	,					
< 2	11>		3	522													
< 2	12>		Ð	NA													
< 2	13>		A.	rabi	ರೆಂಬಕ.	is t	taii.	X1281 (coty	Abc (Colu	moia					
<2	30 >																
<2	21>	•	CI)S													
€27	22>		(:	100)	1	3342	:)										
<22	?3>		At	:NSH:	ful	1-16	nati	LCDI	ia ai	ಗೆ ನ	edua	ad =.	.			se end	
			Pc	y A.De	ptic	le	_			50	. 4661		2 A M C 1	ice (or th	e enc	ಂಡಲಡ
<40)ů>		18	:													
ccs	aagt	esag	c ರ್ಷ (rcgaa	aa t	tggc	aacc	c as	gtto	grca	ı tag	eca:	gae	Caco	;acce	.r.c	60
					ga g									_	,		99
								•									93
Wet	gly	Lye	Gln	Lys	Gln	Gln	Thr	Fla	CCC Cor	cgt	etc	: ttc	get	- 000	aaa Lys		147
1				5	•				10	ALG	2316	Pile	: Ala	Pro 15			
ccc	aaa	tcc	ccg	act	CSC	QAA	CCM	3 a t	225						aca		
510	Lys	Ser	Pro	Thr	His	Glu	Pro	Asn	Pro	VAI	966	Glu Glu	CCS	- Eca	aca Thr		195
			20					25			71.0	0.4	30	261	inr		
ccg	CCE	CCO	aao	ata	tee	~~~			_							`	
Pro	Pro	Pro	LVS	Ile	tee Ser	312	The	gca v.l	ICC	tte	tct	cct	TOO	aag	cgt		243
		35	• -		55.	M. H	40	447	Ser	Pne	ser	2ro 45	Ser	Lye	Arg		
aag	ctt	ote	tee	UMC.	₩	~+~											
Lys	Leu	Leu	Ser	GEÁ	cac His	Leu	Ala	gcc 31=	aca	tca	CCC	233	223	CCL	aaa		291
	50					55	~-a	wra	wrd	\$¢ľ	Pro 6Q	rys	Lys	6x0	Lya		
ctt	tet	CCT	CAC	9.00	<i>_</i>			_ •			_		•				
Leu	Ser	Pro	Hia	The	caa Gln	aac aa-	CCA Dw-	gta	ccc	gat	ccc	aat	tta	cac	caa		339
65				****	Gln 70	Nen	£10	1Ev	Pro	Хвр	Pro	Asn	Leu	Kis	Gln		
-					. 0					75					80		

aga Arg	trt Phe	Leu	G j u Gag	aga Arg 35	ttt Phe	ctg Leu	gaə Clu	Sto	Cog Ser 90	ceg Pro	gag Glu	gaa Glu	tat Tyr	gtt Val 95	Pro	31	87
gaa Glu	acg Thr	t.ca Ser	cca Ser 100	tçg Ser	agg Ar g	aaa Lys	tac Tyr	aca Thr 105	сса Рго	ttg Leu	gaa Glu	cag Gln	csa Gln 110	gtg Val	gtg Val	4.	35
gag Glu	cta Leu	aag Lys 115	agc Ser	Lys	tac Tyr	Pro	gat Asp 120	gtg Val	gtt Val	ttg Leu	atg Met	gtg Val 125	gaz Glu	gtt Val	ggt Gly	46	₿3
tac Tyr	130 yrg agg	tac Tyr	aga Arg	ttc Phe	trc Pbe	934 Gly 135	gaa Glu	yeb gac	Y)a dcâ	gag Glu	atc Ile 140	gca Ala	gca Ala	agc Arg	gtg Val	51	3 2
ttg Leu 145	ggt Gly	att	tac Tyr	gct Ala	cat His 150	atg Met	gat Asp	cac His	аас Абл	Ctc Phe 155	atg Met	acg	νј≖ Зса	agt Ser	gtg Val 160	51	79
cca Cca	aca The	ttt Phu	yrg	ttg Low 165	aac Asn	ttc Phe	cat His	güg Val	aga Arg 1.70	aga Arq	ctg Lev	gtg Val	aat Asn	gca Ala 175	gga Gly	62	2 7 ª
						aag Lys										61	75
						Cly ggc										72	23
						988 Glu 215										77	71
ggt Gly 225	gly	gaa Glu	gaa Glu	GJA	Phe 230	ggt	tca Ser	cag Gln	agt Ser	aat Asn 235	ttc Phe	ttg Leu	gtt Val	tgt Cya	gtt Val 240	81	19
gtg Val	gat Aap	gag Glu	aga Arg	gtt Val 245	aag Lys	tog Ser	gag Glu	aca Thr	tta L a u 250	ggc Gly	tg t Cys	giy Giy	att Ile	gaa Glu 255	atg Met	36	; 7
agt Ser	ttt Phe	gat Asp	gtt Val 260	aga Arg	gtc Val	ggt	grt Val	gtt Val 265	ggc Gly	gtt Val	gaa Glu	att Ile	tcg Ser 270	aca Thr	ggt	91	.5
gaa Glu	gtt Val	gtt Val 275	tat Tyr	gaa Glu	gag Glu	ttc Phe	aat Asn 280	gat Asp	Jas Asn	ttc Pha	atg Met	aga Arg 285	ser	gg a Gly	tta Leu	96	13

									. '•							•
	25	90				29	5	. 71	D MI	a (j](30 n Fe	o Pr T∈	u Le	u Gl	y Gln	
30	5				31	0	. Dy	5 PIJ	a rei	3 V2.	. AI.	a Me	t Al	a Gl	a ect y Pro 320	ioss
				32	5		ı AIG	3 ALS	330	r Leu	l As <u>r</u>	s cy	s Ph	e Se. 33.	-	1107
•			34	10	y (31)	, ASI	. 116	345	Leu	сув	Glu	ı Ly.	9 Ila 35:	e Se: O	c gca r Ala	1155
		35	5		h wal	. Буз	360	ing E	Lys	Lou	Clu	36!	a Ala 5	t Glu	akî r	1203
33a	ate	j te	t tg	e tt	g aca	थरट	car	aca	ATP	AFA						•
	370	;	-,	J 22.	,,,	375	niu	1DI.	ile	Mct	380	Met	, Pro) Rie	3 Lev.	1251
act	gtt	Caa	3 BC	c etc	gcc	cta	ang	ttc	tyc	Cat	ore	aaa	CBO	: FF+	75	. 200
385				J Dec	390	12311	inr	Pne	Cys	81s 395	Leu	Lys	: 6),n	Phe	Gly 400	1299
Phe	gaa Glu	yrg	} ato } Ile	ctt ≥ Leu 405	tac Tyr	caa Gln	g17 âââ	gec Ala	tca Ser 410	ttt Pbc	yrg	t.cc Ser	ttg Leu	tca Ser 415	Ser	1347
			420)	ctt Leu	ser	n_a	425	Thr	Leu	Gln	Gln	130	gag Glu	grt Val	1395
	-,-	435	*******	. Set	gat Asp	GIÀ	5er 440	Glu	ser	Gly	Ser	Leu 445	Phe			1443
atg Met	aat Asn 450	tac His	aca Thr	ctt Lau	aca Thr	gra Val 455	tac Tyr	gct Gly	tcc Ser	yrg :	ctt Leu 460	ctt Leu	aga Arg	cac Hís	tgg Trp	1491
465				200	tgc ; Cys . 470	web 1	AIG I	man .	teu :	tle : 175	ser .	Ala	Arg	Leu	A ap 480	1539
got,q Ala.\	gee /al .	tct Ser	ej <i>n</i> gag	att 11e 485	Ser /	yct (Ala (tge (Cys)	Tet (gga t Gly s	ct o	at & ⊰ie S	agt Ser	Ser	tcc Ser 495	cag Gin	1587

WO 99/19492 PCT/EP98/06977

cte	agc Ser	agr Ser	gag Glu SOO	ecg Leu	gct Val	gaa Glu	gaa Glu	ggs Gly Sos	tct Ser	gag	aga Arg	gca Ala	att 1le 510	gta Val	tca Ser	1635
cet Pro	gag Glu	ttt Phe 515	tat Tyr	CEC CEC	gtg Val	ctc Leu	tcc Ser 520	tca Ser	gtc Val	ttg Leu	aca Thr	gct Ala 525	atg Met	tct S er	aga	1633
								aça Thr								1731
								atg Met								1779
								aag Lys								1927
								aut Thr 585								18 7 5
								gac Aap								1923
								cga Arg								1971
								gag Glu								2019
	_	_	_		_			Y sb gat		_		-			_	2067
								gaa Glu 665								2115
								gar Asp								2163
								act Thr								2211

										-							
70	5			•	,	71	.O		a AI	a re	2 AI:	a Th: 5	r Gl	ı Ri	s Le	t geo u Ala 720	ı)
				-	72.	5		P Ac	r se	736) ∍ ref	a Lys	s Sea	r Ph	9 Se 73	_	1
	•		_	740)	, mj	~ ~1	a MI	8 Va. 74:	1 G11	ı Ala	a Let) Ala	750	l Le	n ysb 2 dac	2355
		7	55				• • •	76	o se	. Arg	. Asn	Lys	765	Tyr	· Va	c cgt l Arg	
	77	O			, ,	1	77:	5	K Pro	o vai	GLu	780	Авл	1]e	Gli	g tot 1 Ser	2451
9 90 Gly 785	Cġ Ar	t c. g H	at (ect Pro	gr.a Val	. 009 Let 780		acı ı Thi	i ata File	tta Leu	caa Gln 795	yab Arć	ysv	ttc Phe	gto Va]	eca Pro	2499
	-				805	20	WT.	. Gil	GIA	810	3.A.r.	Cys	Gln	Ile	11e		2547
·			8	20	OL,	914	гуs	261	625	tac Tyr	Ile	Arg	Gln	Val 830	VJS	ยอก	2595
		₽.3	5		,,,,	9111		840	ser	tt: Phe	Val	Pro	Ala 845	Ser	Phe	Ala	2643
rya	ctg Leu 850	ca Hi	a V	tg (al)	ctt . Leu	gat Asp	855 855	gtt Val	ttc Phe	açt Thr	arg	atg Met 860	ggt (Gly)	gct Ala	tca Ser	gac Asp	2691
965					-	870	≎e1.	THE	Pne		GIU (675	Glu j	Leu S	3er	GI 75	Ala 880	2739
tca Ser	cac His	at:	at e fl	•	iga irg :	acc Thr	сув	tct Ser	tct Ser	egt 1 Arg 8 B90	tog (Ser 1	ett (Leu V	gtt a Val 1	la :	tta Leu 895	gat Asp	2787
g ag (ctt Leu	gga Gly	Ax So	-	ly :	act Thr	agc Ser	THE	cac Kis . 905	gac g Aap (3gt ⊆	jta g /al /	lla I	tt q le 2	jec Na	tat Tyr	2835

														ctt Leu		2893
														cca Pro		2931
														gat Asp		2979
														gtg Val 975		3027
														gcc Ala		3075
						Arg					Mec			aaa Lys		3123
Glu					Ala					Thr				gaa Glu		3171
	gly			Glu					Glu					gct Ala		3219
-	_	_	Бре	_	-			Phe	-				${\tt Gl}{\tt u}$	gac Asp 1055		3267
		Ala					Lys					Ile		GJY ggc		3315
	Arg	cta Leu 1075				Cys			tgai	:ttaa	atc 1		catta	at		3362
agca	actg	jca a	iggt (ttga	gt Ca	tct g	gttag	ş ttg	gegti	ecta	acti	atgi	tgt a	ettag	gtataa	3422
											tgca	agtad	cat o	ettt	cteca	3482
cccó	cata	aaa a	aaaa	aaaa	la a	18445	3888	8 88	aaaa	SEGE					**	3522

<210> 19
<211> 1081
<312> PRT

۷:	213.	•			ትድላይነ	dops	is t	hali	ana	ecot	· vone	CA1	mbia			
e 3	323:	•		1	Polyp	epti	de M	EH3	. Dayer	200	.ype	COLU	mbia			
< 4	000	•		1	L9											
M∈	t C	Зy	Ly	'S (6)	ln Ly	s 01 5	n Gl	aT a	r II	e Se	r Ar	g bp	e Ph	e Ala	a Pro 19) Lys
				•	. •				2	5				3 (•	Thr
Pr	O P	ro	Pr 3	о Г у 5	rs Il	e Se	r Al	a Th 4	r Va	l Se	r Ph	e Se:	r Pro		. Lya	Arg
							3	>				60)			Lys
						r	,				79	5			His	80
					3.	,				.90	0				Val 95	
				401	•				105	•				110		
								120	•				125		Val	_
		_					135			•		140			Arg	
						130					155				Ser	160
					100					170					Ala 175	_
Tyr	Ly	\$ 3	[le	Gly 180	Val	Val	Lys	Gln	Thr 185	GJu	Thr	Ala	Ala	Ile 190	Lys	Ser
Kis	G1;	, , I	la .95	aa4	Arg	Thr	Gly	Pro 200	Phe	Phe	Arg	Gly	Leu 205	Ser	Ala	Lev
Tyr	7h1	: L	ys,	Ala	Thr	Leu	Gኒս 215	Ala	Ala	Głu	Asp	11e 220	Ser	Gly	Gly	Cys
Gly 225	Gly	G	lu	Glu	Gly	Phe 230	Gly	Ser	Gln	5er	Asn 235	Phe	Leu i	Val (Cys :	Val 240
Val	Asp	G	lu	Arg	Val 245	LyL	Ser	Glu	Thr	Leu 250	Gly	Сув	GJY :		31u k 355	let

ser	Phe	Asp	360 A9T	Arg	Val	Gly	Va1	Val 265	Gly	Val	Glu	Ile	5er 270	Thr	Gly
Gle	Val	Val 275	Τγε	Glu	Glu	Ph≑	Asn 230	Asp	Asn	Phe	Met	Arg 285	Ser	gĵy	Leu
Çlu	Ala 290	Val	Ile	Leu	Sor	Leu 295	ser	Pro	Ala	Glu	300	Leu	Ľeu	Gły	nte
9ro 305	Гел	Ser	G1u	Gln	Thr 310	Glи	Lys	Phe	Leu	Val 315	Ala	Met	Ala	Gly	Pro 320
Thr	ser	ΩeA		Arg 325	Val	Glu	УLД	Ala	8er 330	Leu	Yeb	Суа	Phe	Ser 335	Asn
Gly	Asn	Ala	Val 340	Asp	Glu	Val	Ile	Ser 345	Leu	Суз	Glu	ГÀЗ	11e 350	ser	A]a
Gly	Asn	1.eu 355	Glu	Авр	Asp	ГÀВ	G1u 3&0	Met	Lys	Leu	Сĵл	Ala 365	Ala	Glu	ГÀВ
Cly	Met 370	Ser	Cys	Len	Thr	Val 375	Eis	Tiar	Lle	Met	300 ysn	Met	Bra	Bis	Leu
Thr 385	Val	Gln	Ala	í.eu	Ala 390	Lcu	Thr	Phe	СÀЗ	395	Leu	Lys	Gln	Phe	Gly 40 0
Phe	Сји	Arg	Ile	Leu 405	2AL	Gln	СĵÀ	Ala	Ser 410	Phe	Arg	ŞÇY	Leu	Ser 415	Şer
			420					425					Lcu 430		
Val	Lys	Двп 435	пвA	Ser	geA	Gly	Ser 440	Glu	Бer	θĵλ	Ser	Leu 445	Phe	Hig	Asn
Met	Asn 450	His	Thr	Leu	Thr	Val 455	Tyr	Gly	Ser	Arg	Leu 460	Leu	Arg	H3.€	Trp
Val 465	Thr	His	Pro	Leu	Суя 470	Asp	Arg	Ава	Leu	11e 475	Ser	Ala	Arg	Leu	Asp 480
aíA	Val	Ser	Glu	1]e 485	Şer	Ala	Суз	Met	Gly 490		His	Ser	Ser	Ser 495	Gln
Lev	Ser	Ser	Glu SOO		Val	Glu	Ģlu	Gly 505		Glu	Arg	Ala	Ile 510	Va 1	ser
Pro	Glu	9he 51 5	-	, Гел	Val	Leu	5er 520		Val	Leu	Thr	Ala 525	Met	Şer	Arg
Ser	Ser 530	_	Ile	G)n	Arg	G1y		Thr	Arg	lle	Pha 540		Arg	Thr	Ala

- Lys Ala Thr Glu Phe He Ala Val Met Glu Ala He Leu Leu Ala Gly
 545 550 555 560
- Lys Glo Ile Glm Arg Leu Gly Ile Lys Gln Asp Ser Glu Met Arg Ser 565 570 575
- Met Gln Ser Ala Thr Val Arg Ser Thr Leu Leu Arg Lys Leu Ile Ser 580 590
- Val 11e Ser Ser Pro Val Val Val Asp Asn Ala Gly Lys Leu Leu Ser 595 600 605
- Ala Leu Asn Lys Glu Ala Ala Val Arg Gly Asp Leu Lau Asp Ila Leu 610 620
- Ile Thr Ser Ser Asp Gln Phe Pro Glu Leu Ala Glu Ala Arg Gln Ala 625 630 635 640
- Val Leu Val Ile Arg Giu Lys Lou Asp Ser Ser Ile Ala Ser Phe Arg 645 650 655
- Lys Lys Leu Ala Ile Arg Asn Leu Glu Phe Leu Gln Val Ser Gly Ile 660 665 670
- The His Leu Fle Glu Leu Pro Val Asp Ser Lys Val Pro His Asm Trp 675 680 685
- Val Lys Val Asn Ser Thr Lys Lys Thr Ile Arg Tyr His Pro Pro Glu 690 695 700
- The Val Ala Gly Leu Asp Glu Leu Ala Leu Ala Thr Glu His Leu Ala
 705 710 715 720
- Ile Val Asn Arg Ala Ser Trp Asp Ser Phe Lou Lys Ser Phe Ser Arg
 725 730 735
- Tyr Tyr Thr Asp Phe Lys Ala Ala Val Gln Ala Leu Ala Ala Leu Asp
 740 745 750
- Cys Leu His Ser Leu Ser Thr Leu Ser Arg Asn Lys Asn Tyr Val Arg 755 760 765
- Pro Glu Phe Val Asp Asp Cys Glu Pro Val Glu Ile Asn Ile Gln Ser 770 780
- Gly Arg His Pro Val Leu Glu Thr Ile Leu Glu Asp Asn Phe Val Pro
 785 790 795 800
- Asn Asp Thr Ile Leu His Ala Glu Gly Glu Tyr Cys Gln Ile Ile Thr 805 810 815
- Gly Pro Asn Met Gly Gly Lys Se: Cya Tyr Ile Arg Gln Val Ala Leu 820 825 830

Ile	Ser	Ile	Mət	Ala	Gln	Val	GLY	ser	Phe	Val	Pro	Ala	Ser	કેપ્રસ	Ala
		835					849					945			

- Lys Leu His Val Leu Asp Gly Val Phe Thr Arg Met Gly Ala Ser Asp 850 855 860
- Ser Ile Gln His Gly Arg Ser Thr Phe Leu Glu Glu Leu Ser Glu Ala 865 870 875 380
- Ser His Ile Ile Arg Thr Cys Ser Ser Arg Ser Leu Val Ile Leu Asp 885 890 895
- Glu Leu Gly Arg Gly Thr Ser Thr His Asp Gly Val Ala Ile Ala Tyr 300 905 910
- Ala Thr Leu Glo His Leu Leu Ala Glu Lys Arg Cys Leu Val Leu Phe 915 920 925
- val Thr His Tyr Pro Glu Ile Ala Glu Ile Ser Amn Gly Phe Pro Gly 930 940
- Ser Val Gly Thr Tyr His Val Ser Tyr Leu Thr Leu Gin Lya Asp Lyc 945 950 955 960
- Gly Ser Tyr Asp Bis Asp Asp Val Thr Tyr Leu Tyr Lys Leu Val Arg 965 970 975
- Gly Leu Cys Ser Arg Ser Phe Gly Phe Lys Val Ala Gln Leu Ala Gln
 985 990
- Ile Pro Pro Ser Cys Ile Arg Arg Ala Ile Ser Met Ala Ala Lys Leu 995 1000 1005
- Glu Ala Glu Val Arg Ala Arg Glu Arg Ash Thr Arg Met Gly Glu Pro 1010 1015 1020
- Glu Gly His Glu Glu Pro Arg Gly Ala Glu Glu Ser lle Ser Ala Leu 1025 1030 1035 1040
- Gly Asp Leu Phe Ala Asp Leu Lye Phe Ala Leu Ser Glu Glu Asp Pro 1045 1050 1055
- Trp Lys Ala Phe Glu Phe Leu Lys His Ala Trp Lys Ile Ala Gly Lys
 1060 1065 1070

}

- 11s Arg Leu Lys Pro Thr Cys Ser Phe 1075 1080
- <210> 20
- <211> 24
- <212> DNA
- <213> Artificial sequence

<230>		·	
<223>	MSB6 specific primer 638 for PC thaliana ecotype Columbia	R using cDNA of Arabidopsi	5
<400>	20		
cototaco.	ag gegacgaaaa accg		24
<210>	21		
<211>	28		
<212>	DNA		
<2135	Artificial sequence	•	
<320>			
<223>	Primer S81 for PCR using cDNA or Columbia	f Arabidopsis thaliana ecor	ype
<400>	21	•	
cgtcgeett	it aggatocoot tootreae		28
<210>	22		
<211>	30		
<212>	DNA		
<213>	Artificial sequence		
<220>			
<223>	MSH6 specific primer 5823 (or Po thaliana ecotype Columbia	R using cONA of Arabidopsi	5
<400>	22		
gcttggcgc	a tetaatagaa teatgaragg		30
	-	,	
<210>	23		
<211>	24		
<212>	DNA		
<213>	Artificial sequence	•	
<220>	•		
<223>	MSR6 specific primer 637 for PCR thaliana ecotype Columbia	using cDNA of Arabidopsis	
<400×	23		
gacagegee;	a gttcttcaga atgc	· · · · · · · · · · · · · · · · · · ·	24
(210>	24		
211>	33		
212>	DNA		

PCT/EP98/06977

<213>	Artificial sequence	
<230×		
4223>	MSR6 specific primer 188 for FCR using cDNA of Arabidops:	is
	thaliana ecotype Columbia	
<400>	24	•
scareagase	gcagcgreag agategaitt tgt	33
<210>	25	
<211>	27	
<212>	ONA	
<213>	Artificial sequence	
(113)	withing seducates	
<220>		
<223>	MSH6 specific primer SB3 for PCR using cDNA of Arabidops	ai
	thaliana ecotype Columbia	
<400>	`25	
	•	
ogotatetat	ggetgetteg aattgag	27
<210>	26	
<211>	1385	
<212>	CNA .	
<213>	Arsbidopsis thaliana ecotype Columbia	
<223>	Clone 43	
<400>	26	
25020000		
cccgggacge	agogocagag atogattitg tottletter aaaaareeae ggeggegast	60
argaaggqtt	tggtttccgg cgatgctgct agcggcgggg gcggcagcgg aggaccacga	120
TELANTOLOA	aggaagggga tgctaaaggc gacgcttctg tacgttttgc tgtttcgaaa	180
	-5555554-4550 5405-000-05 PEO-1000-50 030-000-34E4	100
tetgregatg	aggitagagg aacggatact ccaccggaga aggiteegeg tegigteetg	240
ccgtctggat	ttaagergge tgaateegre ggtgatgett egteeetgtt etceaatatt	3 D C
atgeataagt	ttgtaaaagt cgatgatcga gattgttttg gagagaggag ccgagaagac	360
at tate come		420
gregiocogo	tgaatgatto atototatgt atgaaggroa atgatgttat tootoaattt	424
cgtcccaata	atggtaaaac tcaagaaaga aaccatgett ttagttteag tgggagaget	480
gascttagat	cagtagaaga tataggagta gatggcgatg ttcctggtcc agaaacacca	540
sanat neat c	CRESTSETTE LEGETTSESS SHEWFFERS ANDERWARD STREET	600
2324-3-4-4	cacgtgcttc tegettgaag ogagttetgg aggatgaaat gaettttaag	,,,,,
gaggataagg	tteetgtatt ggactetaae amaaggetga aaatgeteea ggateeggtt	660

cgeggagag	a agaaagaagi	e aaacgaagg	a accadattt	g aatggcet g:	a gtottotoga	720
atcagggatq	g ccaatagaag	3 acgtectga	t gateeretti	t acgatagaa	a gaccttacac	780
ataccaccto	g atgicticas	a gaaaatgtc	c gearcacaa:	agcaatatt	gagtgttaag	840
agtgaatata	tggacattgt	gettteett	t a as gtgggg	a aattttatga	getgtatgag	900
ctagatgcgg	J aattaggtea	caaggageti	gactggaaga	ı tgaccatgaç	t rāārārāāāās	960
asatgeagae	: aggttggtat	ctctgaaag	gggatagatç	aggcagtgca	aaagctatta	1020
gctcgtggat	: atamagttgg	acgaatcgaq	i cagotagaaa	catetgaces	agcaaaagco	1080
agaggtgcta	aractataat	cccaaggaag	; ctagttcagg	tattaactec	: atcaacagca	1140
agcgagggaa	acatogggoo	tgatgccqtc	: catctlettg	r ctacasasgs	gatcassatg	1200
gagctacasa	agtgttcaac	tgtgtatggs	tttgettttg	ttgactgtgc	tgeettgagg	1260
ttttgggttg	ggtoratoag	ngangangny	teatgtgetg	ccerragage	gttattgatg	1320
caggeetete	caaaqgaagt	gttatatatgac	agcaaagggc	tatesagaga	agcacaaaag	1330
ассскавада	aatatacgtt	gacagggrcc	acggrggcac	agttggctcc	agtaccacaa	1440
gtaatggggg	atacagatgo	tgctggagtt	agazatataa	tagaatotaa	cggatacttt	1500
aaaggecett	cegaatcatg	gaactgtgct	gttgatggtc	taaatgaatg	tgatgttgcc	1560
				taaagctaga		1620
aagcatgggg	atatttttcc	ataccaagti	tacaggggtt	gcotcagaat	cgatggccag	1580
acgatggtaa	arcttgagat	atttaacaac	agctgtgatg	gtggtecsta	agggaccttg	1740
tacaaatate	ttgatametg	tgttagteca	actggtaage	gactottaag	gaattggatc	1800
tgccatccac	tesaagatgt	agaaagcatc	aatsaacggc	ttgatgtagt	tgaagaattc	1860
acggcaaact	cagaaagtat	gcadatcact	ggccagtatc	tocacaaact	tccagactta	1920
gaaagactgc	toggacgcat	caagtotago	gttegateat	cagcototgt	gttgcctgct	1980
ttotgggga	aaaaagtgct	gaaacaacga	gttaaagcat	ttgggcasat	tgtgaaaggg	2040
tragaagtg	gaattgatcc	gttgergger	ctacagaagg	aaccaaacat	gatgagtttg	2100
tttataaac	totgtaaact	toctacatta	gtaggaaaaa	gegggetaga	gttatttctt	2160
ctcaattcg	aagcagccat	agatagçç				2195

7500

1260

<210>	27					
<212>	1385					
<212>	DNA					
<213>	Arabidons	sie thaliana	a acotime Cr	al umbai a		
<223>	Clone 62		- ceoolbe co	71 U.I.C.1 A		
~ 667.	315/19 02					
<400>	27					
catcageete	tgtgttgcct	getettetgg	ggaaaaaagt	getgaaacaa	dgagttakag	60
catttgggca	aattgtgaaa	gggttcagaa	gtggaattga	totgttgttg	gctctacaga	120
aggaatcaaa	tatgatgagt	ttgctttata	aactctgtma	acttectata	ttagtaggaa	130
aaageggget	agagttattt	ctttctcaat	tognageage	catagatage	gactttccaa	240
attatcagaa	ccaagatgtg	acagatgaaa	acgetgaaae	teteaesata	cttatçgaac	300
tttttatega	aagagcaact	caatggtetg	aggicatica	caccataagc	cgcctagatg	360
tortgagato	ctuegoaard	gcagcaagtc	tototgetgg	aagcatggcc	aggeetgtra	420
					atacttazaa	480
					gttccgaatg	540
				teatectegg	-	600
					ctggccgtta	660
				cgaaatetee		720
				aggagagagt.		780
				tcaggattca		840
				cgccattgca		900
					taccaccete	960
				acacatggct		1020
aatcaagatc	tgattatcaa	ccacgtggtt	gtgatcaaga	cctagtgttc	ttgtaccgtt	1080

agreattigt gigitatitet egagtegeee acaacaatge ecceattige gaagatgaet 1320 aegacketti gittigetta tigeatgaga teaaateete thacigigit eccaaataac 1390

taaccqaqqq agcttgtcot gagaqctacq gacttcaagt ggcactcatq gctgqaatac

caaaccaagt ggttgaaaca gcatcaggtg ctgctcaagc catgaagaga tcaattgggg

addacticaa gicaagigag ciaagarcig agiteteaag-trigcatgaa gactggetca

```
ceggg
                                                                         1385
 <210×
               28
 <211>
               34
 <212>
              DNA
 <213×
              Arrificial sequence
 <220>
              MSH6 specific primer 288 for PCR using cDNA of Arabidopsis
 <223>
              thaliana ecotype Columbia
 <400>
              26
 atorogggtt atttgggman anagtaagag gatt
                                                                          34
 <210>
              29
 <211>
              27
 <212>
              DNA
 <213>
              Artificial sequence
 <220×
             MSH6 specific primer SB2 for PCR using cDNA of Arabidopsis
 <233>
              thaliana ecotype Columbia
<400>
              29
gegitegate atcagectet gigitge
                                                                          27
<210>
             30
<211>
             3606
<212>
<213>
             Arabidopsis thaliana ecotype Columbia
<220>
<221>
             CDS
<222>
             (142) . . . . (3468)
             AtMSH6 full-length command deduced sequence of the encoded
<223>
             polypeptide
<400>
             30
assagttgag ccctgaggag tatcgtttcc gccatttcta cgacgcaagg cgaaaatttt
                                                                        60 .
tggcgccaat ctttcccccc tttcgaatte tctcagctea aaacatcgtt teteteteae
                                                                       120
teteteteae aattecaasa a atg cag ege cag aga teg att tig tet tie
                                                                       171
                        Met Gln Arg Gln Arg Ser Ile Leu Ser Phe
```

								Thr								219
got Ala	gct Ala	agc Ser	30 CJÀ 33c	ggg Gly	GJY ÿ g c	eja Gac	agc Ser	35 GLY GB#	Gjy gga	cca Pro	ega Arg	tct Phe	aat Asn 40	gtg Val	aag Arg	257
								tct Ser								315
								gat Asp								363
	-	-	-					aag Lys	_	_	-			-	-	411
			_					atg Mes				-			-	459
-								agc ser 115							ctg Leu	507
								gct Ala								5 5 5
								gaa Glu								603
_		-	_	•		_		gta Val	_	_			_	_		651
-					_			G1y 9 9 9	_			_			_	699
								atg Met 195								747
			_					ctg Leu					-			795

																• • •
CÁB	gga Gly 220	Giu	ГÀ2 493	aaa Uye	Glu Gaa	gta Val 225	aac Asn	gaa Glu	gga Gly	acc	asa Lys 230	ttt Phe	gaa Glu	Trp	ott Leu	943
gag Glu 235	ser	tat Ser	ega Arg	ace Ile	agg Arg 240	gæt Asp	gcc Ala	tss nak	aga Arg	aga Arg 245	yrg	eet Pro	gat Asp	gat Asp	ece Pro 250	891
ctt Leu	tac Tyr	yeb	aga Arg	aag Lys 255	acc Thr	tta Leu	cac Bis	ata Ile	cca Pro 260	cct Pro	gat Asp	gtt Val	ttc Phe	aag Lye 265	aaa Lys	939
atg M et	tct Ser	Sca Sca	tca Ser 270	Gln	aag Lys	Gju Cas	tat Tyr	tgg Trp 275	agt Ser	gtt Val	aag Lys	agt Ser	gaa Glu 280	tat Tyr	atg Met	987
gac qu	att Ils	gig Val 285	ctt Leu	ttc Phe	trt Phe	aaa Lys	gr.g Val 250	61 <i>à</i> 88 8	asa Lys	ttt Phe	tat Tyr	gag Glu 295	ctg Leu	tat Tyr	Glu Gag	1035
cta Leu	gat Asp 300	gcg Ala	gaa Glu	tta Leu	83y 8gt	cac Mis 305	aag Lys	gag Glu	ren	yab d s c	tgg 7rp 310	aag Lya	atg Met	acc ant	atg Met	1093
agt Ser 315	ggt Gly	8°3	GTÅ ââs	Lys	tgc Cys 320	aga Ary	cag Gla	gtu Val	gjy Gjy	ate Ile 325	tct Ser	gaa Glu	agt Ser	era 333	ata Ile 330	1131
gat	gag Glu	gca Ala	grg Val	caa Gln 335	PAB	ctà Leu	tta Leu	gec Ala	cgt Arg 340	gjy gga	tat Tyr	aaa Lys	gtt Val	gga Gly 345	Axa cas	1179
ate Ile	gag Glu	cag Gin	cta Leu 350	gaa Glu	aca Thr	tct Ser	yab Bgc	caa Gln 35 5	gca Ala	aaa Lys	A¦a Ala	yrg	ggt Gly 360	get Ala	aat Asn	1227
act	ata Ile	att Ile 365	6x6 ccs	agg	aag Lya	cta Leu	gtt Val 370	cag Gln	gta Val	tta Leu	act Thr	cca Pro 375	tca 8er	aca Thr	gcá Ala	1275
agc Ser	gag Glu 300	ej aas	aac Abn	atc Ile	G1 y ggg	cct Pro 385	gat Asp	gce Ala	gtc Val	cat His	ctt Leu 390	ctt Leu	get Ala	ata Ile	aaa Lyb	1323
gag Glu 395	atc	aaa Lys	atg Met	g a g Glu	cta Leu 400	caa Gln	aag Lys	tgt Cys	tca Ser	act Thr 405	gtg Val	tat Tyr	gga Gly	ttt Phe	gct Ala 410	1371
ttt Phe	gtt Val	gac Asp	tgc Cya	gct Ala 415	gre	t t g Lev	agg Arg	bµe	tgg Trp 420	gtt Val	gly ggg	tec Ser	atc Ile	agc Ser 425	gat Asp	1419

_	-			-	_			gcg Ala 435		.,	_	_	•			1467
_	_				_			gjå 833			_	_	_		_	1515
_					_	_		9JÀ 8 3 2				_	_	_	_	1563
	_			-	_			aca Thr	_	_	_		_	-		1611
		_						aaa Lys	-			_				1659
_	•	-	_					cgs Cys 515								170 7
								yrg					Азр			1755
_			_					caa Gln	-			Gly				1803
						val		Lev CCC			Spe					1851
					Gly			tac Tyr		Tyr						1899
								agg Arg 595						Pro		1947
			Glu					cgg Arg					Glu			1995
_	_	Asn		_	-	_	ĠζΝ	atc lle				TYI				2043

635	5	a vef	, re/	946	640 640	Leu) Lec	2 Gly	/ Arg	;	: Lye	Se:	r Sei	r Val	cga L Arg 650	2091
t ca Sær	tes Sei	s gro r Ala	tet Ser	gtg Val	ال جيا	ect Pro	Ala gct	ctt Leu	ctg Leu 660	r GJA	, aaa Lys	aaa Lys	agte Val	ctg Lev 669	y aaa 1 Lys 5	2139
caa Gln	oga Arg	l Asj	400 Lye 670	мја	ttt Phe	Gly 9 9 9	Caa Gln	att ile 675	Val	Lys Lys	. 617 338	tto Phe	aga Arç	Se:	99a 61y	2187
att Ile	gat Asp	ctg Leu 685	וופיע	ttg Leu	gct Ala	Cta Leu	Gln 690	Lys	gaa Glu	tca Ger	aat Asn	atg Met	Met	agt Ser	ttg Leu	2235
cet	tat Tyr 700	DAR	ct c Leu	tg t Cys	aaa Lys	ctt Leu 705	cct Pro	ata Ile	tta Leu	gta Val	gga Gly 71n	Lys	ago Ser	. Gly saa	cta Leu	2283
gag Glu 715	LZU	ttt Phe	ctt Leu	tct Ser	caa Gln 720	t t c Phe	gaa Glu	gca Ala	yla	ata Ile 725	ysp	agc Ser	ysp	ttt Phe	cca Pro 730	2331
ast	tat Tyr	Gla	aac Asn	caa Glp 735	gat Asp	gtg Val	aca Thr	gat gat	gaa Glu 740	aac Aan	get Ala	gaa Glu	act Thr	ctc Leu 745	aca Thr	2379
ata Ile	ctt Leu	atc Ile	gaa Glu 750	ctt Lev	ttt Phe	ate	gaa Glu	aga Arg 755	gca Ala	act Thr	caa Gln	tgg Trp	tet Ser 760	Glu	gte Val	2427
att Ile	cac His	acc Thr 765	ata Ile	agc Ser	tgc Cys	cta Lev	gat Asp 770	gtc Va).	ctg Leu	aga Ang	tct Ser	ttt Phe 775	gca Ala	atc Ile	gra Ala	2475
gca	agt Ser 780	ctc Leu	tct Ser	gct Ala	gga Gly	age Ber 785	MOL	gcc Ala	agg	cct Pro	get Val 790	att .Ile	tct. Phe	ecc Pro	gea Glu	2523
tca Ser 79\$	GjA Ĝ95	gct Ala	aca Thr	gat	cag Gln 800	aat Asn	ca g Gin	aaa Lys	aca Thr	aaa Lye 805	G1 y 333	cca Pyd	ata Ile	ctt Leu	aaa Lys 810	2571
ate I le	caa Gln	gga Gly	cta Leu	tgg Trp 815	çat His	cca Pro	ttt Phe	aia	gtt Val 820	gca Ala	gcc	gat Asp	gly agc	caa Gln B25	ttg leu	2619
est Pro	gtt Val	crg Pro	aat Asn 830	gat Asp	ata Ile	ctc Leu	Leu	ggc Gly 835	gag Glu	get Ala	aga Arg	aga A rg	agc Ser 840	agt Ser	gg¢	2667

											cca Pro					2715
961	115	345	210	Ag	267	Leu	950	eeu	4611	GIY	PIO	855	MEE	GIY	GIY	
											gtt Val					2763
-,,	960					865		-,-			870	2,4				
ett	aac	tac	tac	ata	CCD	+	020	tet	t a c	ma =	ato	٧,,~	et e	at a	an t	2011
											Ile					2811
875		-	-		880	•				B85					890	
act	atc	tte	aca	ಸರ್	ctt	age	gca	nae	gat.	вда	atc	ato	ara	cas	gag.	2859
											Ile					2017
				895					900					905		
agc	acc	ctt	ttg	gta	gaa	tgc	act	gag	aca	gcg	tca	gtt	ctt	cag	aat	2907
			Leu					Glu			Şer		Leu			
			910		•			915					320			
											crg					2955°
λla	Thr		Asp	Ser	Leu	٧a ٤		Leu	Asp	ิดวิน	i.eu		yrd	gjy	Thr	est.
		925					930					935				÷
											gtt					3003
Şer	Thr	Phe	Asp	Gly	Tyr	Ala 945	IJέ	Ala	Tyr	rss	7al 950	БрБ	Arg	His	Leu	
	340					343					750					
											aca					3051
955	Glu	Lys	Val	Gln	960	Arg	Met	Leu	Phe	A1a	The	His	Tyr	His	Pro 970	
,,,,					300					703					3,0	
		_	_						-	•	300	_			_	3099
ľ÷ፖ	Thr	rhs	GIU	975	Ala	Ser	His	Pro	BB0	V&1	Thr	ser	Lye	985	Met	
				, . .					300					,,,,		
_	-	_				_		_			CC St	_		_	_	3147
AT 3	Cys	Ala	990	Lys	ser	vtd	ser	995	тут	GIV	5 L O		7000 Gta	Суб	Asp	
								3.3					3 000			
											gga					3195
GIN	_	1005	VAL	517€	rea	_	Arg 1016	DER	Inr	GIB	Gly	ALA Sigi	Cys	PLO	Giu	
						·										
											ata Ile					3243
	102D	GtÅ	₽≅u	ATIF		1025	LEU	MEC	WIG		1030	PIO	ASI	Olu	V&1	
											aag Lys					3291
103		1 115	A.I.d.		G19 1040		wig			7045	PAA	wrA	Ser		GLY 1050	
															-	

			,	105	1 5e 5	t ga r Gl	V L€	u Ar	g Se 106	r Gl O	u Ph	e Se	x Şe	106	u Xi 5	8	3339
			107	0	p 74	a tt: F Lei	u va	107	у II 5	8 3e	r Ar	g Va	1 Al 108	а Ні 0	s as	n	3387
		108	5	9 01	, 41	n yai g dal	109	0 D 13	r A9	o Tr	r Le	u Ph 109	e Cy 5	8 fe	a tg u tr	b a	3435
,,,,	110	0	= Dy	3 5 6:	c se	t tad r Tyr 1105	г Су: 5	9 Va:	l Pr	o Lya	.						3478
						ctcgt											3538
384	aat	gctt	atai	tatea	laa a	iaatit	به بان	ני ט:	cegai	ctaar	ર હાસ	13336	aaa	3333	aaaa	AAA	3598
aaa	laaaa	Ræ															3606
	1> .2> .3>		R9 A1	.09 T æðid	'opsi	s th	alia	na e	coty	jb⇔ c	'olum	bia					
< 40			31		pero	e fis.	πь.										
					_												
1	911	ura	. 911	Arg 5	ser	ile	гел	. Ser	jo Jye		Gln	Lys	Pro	Thr 15			
Ala	Thr	Thr	Lys 20	gly	Гфл	Val	Ser	Gly 25	, YED	Ala	Ala	Ser	Gly 30		Gly		
Gly	Ser	Gly 35	Gly	Pro	Arg	Pħę	A\$D 40	Val	Arg	СJЛ	gly	Лар 45	λla	Lys	Gly		
Asp	Ala 50	Ser	Va1	Arg	Phe	Ala 55	Va L	ser	Lye	Ser	Val 60	Asp	Glu	Val	yrğ		
G1y 65	Thr	qak	Thr	Pro	Pro 70	Glu	Lys	Val	Pro	Arg 75	Arg	Va1	Leu	Pro	9 e r		
Gly	Phe	Lys	Pro	Ala 85	Glu	ser	Ala	Сĵу	Авр 90	Ala	Ser	Ser	Leu	Phe 95	Şer		
Aan	Yle	Met	eiK OO <i>l</i>	Lya	Phe	Val	Lys	Val 105	Asp	Asp	Arg	Ар	Cys 110	Ser	gly		
Glu	Arg	ser 115	Axg	01 <i>n</i>	Asp	val	Val 120	Pro	Leu	Asn	qeA	Ser 125	Ser	Leu	Суз		

Met	L/S 130	Ala	asa	Asq	Val	11 0 135	Pro	Gln	Phe	yrg	Ser 140	Asn	Aan	ĠĮÅ	Гуз
Th¥ 145	6 1 8	Glu	Arg	Asn	Ні з 150	Ala	Phe	8er	Phe	Ser 195	Gly	ЯгЗ	Ala	Glu	ren Ten
ртя	Ser	val	Glu	Asp 165	Ile	Gly	Val	Asp	Gly 170	Двр	Val	PTO	GJA	Pro 175	Glu
Thr	Pro	Gly	Met 180	Arg	Pro	Arg	Ala	8er 185	Arg	Leu	Ъуs	Arg	Val 190	Leu	G) u
Aep	Glu	Met 195	The	ይがፍ	Lys	Glu	Asp 200	Lys	Val	Pro	Val	Leu 205	дар	Ser	Aan
Lys	Arg 210	Leu	Lys	Met	Гел	Gln 215	Asp	Pro	Val	Cys	Gly 220	Glu	Liye	ГÀВ	Glu
Vа1 225	Asn	Glu	Gly	Thr	Ĺуs 230	Phe	Giu	dr _L	Leu	Clu 235	Scr	5 er	УīЗ	IJe	Arg 240
Asp	Ala	Asn	ard	Arg 245	Arg	Pro	ASP)	Acp	Pro 250		Tyr	ДЗР	Arg	Lys 255	The
Lou	Bis	Ile	250	Pro	Asp	val	Рλе	Lys 265	Lys	Жet	Ser	Ala	Ser 270	Gln	Lys
Glp	Туг	1'rp 275	Ser	Val	Lys	Ser	3ВО С]л	Tyr	Met	Авр	IJe	Va) 285		Phe	Phe
Lys	Val 290	GJY	Lys	Phe	Tyr	Glu 295	Leu	Туг	Glv	Leu	Asp 300		Glu	Гел	GJ?
Нів 305	Ľув	Glu	Leu	Asp	Trp 310	ГАв	Met	Thr	Met	Ser 315		Val	GJA	Lya	330 CAs
Arg	Gln	VA1	G1y	Ile 325		СJл	Ser	Glγ	13e 330		Glu	Ala	Val	Gln 335	Lys
Leu	Leu	Ala	Arg 340	GJγ	Тух	Lys	Val	Gly 345	_	Ile	Głu	Gln	leu 350	Glu	Thi
Ser	Авр	Cln 355	Ala	Lys	Ala	Arg	Gly 360		Àatî	Thr	Ile	11e 365		Arg	Lys
Leu	∀a1 370	Gln	۷al	Leu	Thr	Pro 375		Thr	Ala	, \$er	380		Asto	lje	Gly
Pro 385	•	Ala	Val	His	Leu 390		Ala	lle	Lys	Glu 395		Lys	Met	Glu	Let 400
Gln	Lys	Сув	Ser	Thr 405		Tyr	Gly	Phe	Ala 410		Val	Авр	Суъ	Ala 415	

- Leu Arg Phe Trp val Gly Ser lie Ser Amp Amp Ala Ser Cym Ala Ala 420 425 430
- Leu Gly Ala Leu Leu Met Gln Val Ser Pro Lys Glu Val Leu Tyr Asp
 435 440 445
- Ser Lys Gly Leu Ser Arg Glu Ala Gln Lys Ala Leu Arg Lya Tyr Thr 450 455 460
- Leu Thr Gly Ser Thr Ala Val Oln Leu Ala Pro Val Pro Gln Val Met
 465 470 475 480
- Gly Asp Thr Asp Ala Ala Gly Val Arg Asn Ile Ile Glu Ser Asn Gly 485 490 495
- Tyr Phe Lys Gly Ser Ser Glu Ser Trp Asn Cya Ala Val Asp Gly Lau
 500 505 510
- Asn Glu Cys Asp Val Ala Lou Ser Ala Leu Cly Glu Leu Ile Asn His 515 520 525
- Leu Ser Arg Leo Lys Leu Giu Asp val Leu Lys His Gly Asp Ile Phe 530 535 540
- Pro Tyr Gln Val Tyr Arg Gly Cys Leu Arg Ile Asp Gly Gln Thr Met 550 555 560
- Val Aan Leu Glu Ile Phe Aan Aan Ser Cys Aap Gly Gly Pro Ser Gly 565 \$70 575
- The Leu Tyr Lys Tyr Leu Asp Asn Cys Val Ser Pro The Gly Lyc Arg
 580 585 590
- Leu Leu Arg Asn Tro Ile Cys Bis Pro Leu Lys Asp Val Glu Ser Ile 595 600 605
- Asn Lys Arg Leu Asp Val Val Glu Glu Phe Thr Ala Asn Ser Glu Ser
- Met Gln Ile Thr Gly Gln Tyr Leu His Lys Leu Pro Asp Leu Glu Arg 625 630 635
- Leu Lau Gly Arg Ile Lys Ser Ser Val Arg Ser Ser Ala Ser Val Leu 645 650 655
- Pro Ala Leu Cly Lys Lys Val Leu Lys Gln Arg Val Lys Ala Phe
- Gly Glm Ile Val Lys Gly Phe Arg Ser Gly Il Asp Leu Leu Ala 675 680 585
- Leu Gln Lys Glu Ser Aan Met Net Ser Leu Leu Tyr Lys Leu Cys Lys 690 700

CONTINUE OF STREET

Сеч 705	B ID	Ile	Гев	۷al	710	Lys	Ser	Gly	ГБЛ	61u 715	Leu	БрG	Leu	Ser	G1n 730
БрБ	GĴu	Ala	Ala	Ile 725	Asp	Ser	Asp	Phe	Pro 130	Asn	Tyr	Gln	Asn	Gln 735	Asp
Val	Thx	qeA	Glu 740	æπ	Alu	Glu	Thr	Նeu 745	Thr	Ile	Leu	Ile	Glu 750	Leu	Phe
Ile	Glu	Arg 755	Ala	Thx	Gln	Trp	Ser 760	Glu	Va1	Île	Kia	Thx 765	Ile	Ser	Cys
Гвл	А sp 7 7 0	Vul	Leu	Arg	Ser	₽ħ₽ 775	A: a	lle	s La	Ala	ser 780	Leu	\$er	Ala	Gly
Ser 785	Met	Ala	Arg	Pro	Val 790	lle	Phe	Pro	Glu	Ser 795	Glu	Ala	Thr	Asp	800
Aso	Gln	Lys	Thr	Lys 805	GJλ	Pro	1:e	l.eu	Lys 810	Lle	Gla	GJY	Lev	Trp 815	Nis
Рго	Đœ	Ala	Va). B20	Ala	λla	ysb	21 À	832 GJv	Leu	Pro	٧æ١	Pro	088 088	Asp	Ile
Leu	Leu	Gly 935	Glu	Ala	yrg	Arg	Ser 640	Ser	GJA	Ser	Ile	His B45	Pro	Arg	Ser
ren	Leu 650	ren	Thr	Gly	Pro	Asn 855	Met	Gly	Cly	Lys	Ser 860	Thr	Leu	řen	Arg
Ala 865	Thr	Суѕ	Leu	Ala	Val 870	Ile	₽'ne	Ala	Gln	675	gly	СУэ	туг	Val	880 Pro
Суб	Glu	Ser	Суѕ	Olu BBS	Ile	ser	Leu	Val	Asp 890	Thr	Ile	Phe	Thr	Arg 395	Leu
Gly	Ala	Ser	Asp 900	Arg	Ile	Mel	Thr	Gly 905	Glu	Ser	Thr	Phe	200 910	Val	Glu
Сув	Thr	Glu 915	Thr	Ala	Ser	Val	Leu 920	Gln	Asn	Ala	Thr	Gln 925	Авр	Ser	Lev
Val	11e 930	Leu	Asp	Glu	Leu	935	yrg	Gly	Thr	Ser	Thr 940		Ąsp	Gly	Tyr
Ala 945	Ile	Ala	Тут	Ser	Val 950	Phe	Arg	His	Leu	Val 955	Glu	ГЛВ	Va)	Gln	Cys 960
Arg	Met	Leu	Phe	Ala 965	Thr	Жiв	Tyr	BLK	Pro 970	Leu	Thr	ГÀВ	Glu	Phe 975	Ala
ser	Ķis	Pro	Arg 980	Val	Thr	Ser	Lys	His 985	Met	Ala	ÇAs	Ala	Рћа 990	Lys	Ser

Arg Ser Asp 995	Tyr Gln Pro Arg Gly Cys Asp Gln Asp Leu Val Phe Leu 1000 1005	
Tyr Arg Leu 1010	Thr Glu Gly Ala Cys Pro Glu Sex Tyr Gly Leu Gln Val 1015 1020	
Ala Leu Mot 1025	Als Gly Fie Pro Asn Gln Val Val Glu Thr Ala Ser Gly 1030 1035 1040	
Ala Ala Gln	Ala Met Lys Arg Ser Ile Gly Glo Asn Phe Lys Ser Ser 1045 1050 1055	
Glu Leu Arg	Ser Glo Phe Ser Ser Leu His Glu Asp Trp Leu Lys Ser 1065 1070	
Leu Val Gly 1075	fle Ser Arg Val Ala His Asn Amo Ala Pro Ile Gly Glu 1080 1085	
Asp Asp Tyr 1090	Asp Thr Law Phe Cys Lew Trp His Glu IIe Lys Ser Ser 1095 1100	
Tyr Cys Val 1105	Pro lys	
<210>	32	
<212>	24 DNA	
<213>	Artificial sequence	
<220>		
<223>	Portuged Animov for the analysis of amounts	
- 	Forward primer for FCR amplification of ATRGENEA microsatellite	
<400>	32	
accatgoata o	CTTABACTT CTTG	24
<210>	33	
<211>	22	
<212>	DNA	
<213>	Artificial sequence	
<320>		
<223×	Reverse primer for PCR amplification of ATHGENEA microsatellite	
<400>	33	
200020000	33735555 40	

<2105	34			•
<211>	18			
<212>	cna			
<213>	Artificial sequence			
<220>				
<223>	Forward primer DMCIN-A for PCR	on genemic	DNA of	Arabidocsis
	thaliana ssp. Landsberg eracta	"Ler"		·
<400>	34			
gaagogatat	tqttcgtg			18
<510>	35			
<211>	13	•		
<212>	DNA			
<213>	Artificial sequence		-	
<220>				
c223>	Reverse primer DMCIN-B for PCR thaliana ssp. Landsberg erecta		DNA o	f Arabidopaja
<400>	35			
agattqcg a g	aacattcc			18
<210>	36			
<211>	31			
<212>	DNA			
<213>	Artificial sequence			
	1111111111			
<220>				
<223>	Forward primer DMCIN-1 for PCR thaliana map. Landaberg erecta	•	DNA of	E Arabidopsis
<400>	36			
acgcgtcgac	tcagctatga gattactcgt g			31
<210>	37			
<211>	29			
<212>	DNA			
<213>	Artificial sequence			
<220>				
c223>	Patraga primar MMATU 2 5a- 000		DUB C	S Swahidanai-
	Reverse primer DMCTN-2 for PCR thaliana ssp. Landsberg erecta		DIA O	. Wiggioobs/9
<400>	37			
getetagate	tetequicta agactetet			29

```
<210>
               38
  <211>
               32
  <212>
               DNA
  <2135
               Artificial sequence
 <2205
  د2233ء
              Forward primer DMCIN-3 for PCR on genomic DNA of Arabidopsis
              thaliana esp. Landsberg erecta "Ler"
 <400×
               38
 getetagage tretettaag taagtgattg at
                                                                            32
 <210>
              39
 <211>
              48
 <212%
              DNA
 <2135
              Artificial seguence
 <2205
              Reverse primer DMCIN-4 for PCR on genomic DNA of Arabidopsis
 :223×
              chaliana sap. Landsberg erecta "Ler"
 <400>
              39
 tootoogggs togagagate tecatggitt citeagetet atgaatee
                                                                           48
<210>
              40
<2115
              26
<2125
              DNA
<213>
              Artificial sequence
<220>
<223>
             Forward primer DMC1a for PCR on genomic DNA of Arabidopeis
             thaliana asp. Landsberg erects "Ler"
<400>
             40
acycytrgac gaattrycaa gtygyy
                                                                           26
<210>
             41
<211>
             38
<212>
             DNA
<213>
             Artificial sequence
<220>
<223×
             Reverse primer DMClb for PCR on genomic DNA of Arabidopsis
             thalians ssp. Landsberg erecta "Ler"
<400≥
             41
```

nccatggaga	totocogggt acceatetge ttogaggg	38
<210>	42	
<212>	20	
<212>	DNA	
<213>	Artificial sequence	
<220>	-	
<223>	Forward primer for PCR amplification of ATEAT1 SSLP marker Arabidopsis thaliana subapecies	rir
<400>	42	
gccactgcgt	gaatgatatg	20
<210>	4.3	
<2115	22	
<212>	DNA	
<213>	Artificial sequence	
10237	merrera bedustive	
<220>		
<2235	Reverse primer for PCR amplification of ATEAT1 SSLP marker Arabidopsis thaliana subspecies	ir
<400>	43	
*********	acattaattc cc	
cyaacageca	acattaattu co	22
<210>	44	
<311>	18	
<212>	DNA	
<313>		
4323 3	Artificial sequence	
<220>		
<223>	Forward primer for PCR amplification of MGA63 SSLP marker Arabidopsis thaliana subspecies	in
<400>	44	
aaccaaggca	cagaagcg	18
<210>	45	
<211>		
<211×	15 DNA	
<213>	—-· - -	
49172	Artificial sequence	
<220>		
<223>	Reverse primer for PCR amplification of NGA63 SSLP marker	in

```
<400>
                45
   acceaagiga tegecace
                                                                              18
  <210>
                46
  <211>
                21
  <212>
                DNA
  <213>
               Artificial sequence
  <220>
  <233>
               Forward primer for PCR amplification of NGA248 SSLP marker in
               Arabidopsis thaliana subspecies
  <400>
               46
  tассувасся азасасява<u>я</u> д
                                                                             21
 <210>
               47
 <211×
               22
 <212>
               DNA
 د213×
               Artificial sequence
 <220>
 <223>
              Reverse primer for PCR amplification of NGA248 SSLP marker in
              Arabidopsis thaliana subspecies
 <400>
               47
 totgtatoto ggtgaattot co
                                                                            22
 <210>
              48
 <2115
              22
<212>
              DNA
<213>
              Artificial sequence
<220>
<223>
              Porward primer for PCR amplification of NGA128 SSLP marker in
              Arabidopsis thaliana subspecies
<400>
              48
ggtctgttga cgtcgtaagt cg
                                                                           22
<210×
              49
<211>
             22
<212>
             DNA
<213>
             Artificial sequence
<220>
```

WO 99/19492 PCT/EP98/06977

<2235	Reverse primer for PCR amplification of NGA128 SSLP marker in Arabidopsis thalians subspecies
<400>	49
atottgaaac c	tttagggag gg 22
<210>	50
<211>	22
<212>	DNA
<213>	Artificial sequence
	AIBIIIOIAI GEQUENCE
<220>	
<223>	Forward primer for PCR amplification of NGA280 SSLP marker in
	Arabidopsis thaliana subspecies
<400>	S O
ctgatctcac g	gacaatagt gc 22
	•
<210>	51
<211>	20
<212>	DNA
<213>	
<213>	Artificial sequence
<220>	
<223>	Reverse primer for PCR amplification of NGA280 SSLP marker in
1000-	Arabidopsis thaliana subspecies
	The state of the s
<400>	51
ggotecataa aa	agacacacacacacacacacacacacacacacacacaca
<210>	52
<211>	21
<212>	DNA
<213>	Artificial sequence
	Wiringt sedmence
<220>	
<223>	Forward primer for PCR amplification of MGAlll SSLP marker in
	Arabidopsis thaliana subspecies
<400>	52
ctccagttgg a	agctaaagg g 21
	`
c210>	53
<211>	21
<212>	DNA
<2135	Artificial seguence

<220>	
<223>	Reverse primer for PCR amplification of NGAlll SSLP marker in Arabidopsis chaliana subspecies
<400>	53
tgttttta	g gacaaatgge g
	21
<210>	54
<211>	20
<212>	DNA
<213>	Artificial sequence
<220>	
<233>	Forward primer for PCR amplification of NGA168 SSLP marker in Arabidopsis thaliana subspecies
<400>	5-4
ccttcacato	• Caaaacccac 20
.240	
<210><211>	55
<212>	20
<213>	DNA
-0237	Artificial sequence
<220>	
<223>	Reverse primer for PCR amplification of NGALES ESLP marker in Arabidopais thaliana subspecies
<400>	55
gcatataecc	acaaccagaa 20
<210>	
<211>	5 <i>6</i> - 20
<212>	
<213>	DNA
	Artificíal sequence
<220>	
<223>	Forward primer for PCR amplification of NGA1126 SSLP marker in Arabidopsis thaliana subspecies
<400>	56
cgctacgctt	ttoggtasag 20

PCT/EP98/06977

WO 99/19492

<210>	57
<211>	20
<212>	UNA
<213b	Artificial sequence
<220>	
€2235	Reverse primer for PCR amplification of NGA1126 SSLP marker in Arabidopsis chaliana subspecies
<4Q0>	57
අරස්ථස්වුවරරා	agteacaace 20
<210>	
<211>	5 8
<211>	20 DNA
<213>	Artificial sequence
<8T3>	Altiticisi sequence
<2 20 >	• · · · · · · · · · · · · · · · · · · ·
<223>	Porward primer for PCR amplification of NGA361 SSLP marker in Arabidopsis Chulinna subspecies
<400>	58
aaagagatg	gaatttggar 20
<310≻	59
<211>	23
<212>	DNA
<213>	Artificial sequence
<220>	
<223>	Reverse primer for PCR amplification of NGA361 SSLP marker in Arabidopsis thaliana subspecies
<400>	59
acatatcaa	: atattaaagt agc 23
<210>	60
<211>	18
<212>	DNA
<213>	Artificial sequence
<22 0 >	
<323>	Porward primer for PCR amplification of NGA168 SSLP marker in Arabidopsis thaliana subspeci s
<400>	60

ceatchacts	cactgeco	
,	18	
<210>	61	
<211>	32	
<212>	DNA	
<213>		
C2135	Artificial sequence	
<220>		
<223>	Reverse primer for FCR amplification of NGA168 SSLP marker in Arabidopsis thaliana subspecies	1
<400>	61	
gaggacatgt	ataggageet eg	
<210>	62	
<211>	20	
<212×	DNA	
<21,3>	Artificial sequence	
<220>		
<223>	Powered and and the second	
	Porward primer for PCR amplification of Athbroz SSLP marker in Arabidopsis chaliana subspecies	
<400>	62	
tgacctcctc	ttecatggag 20	
210>	63	
:211>	22	
(212>	DNA	
:213>	Artificial sequence	
:320>		
:223>		
	Reverse primer for PCR amplification of AthBIO2 SSLP marker in Arabidopsis thatians subspecies	
400>	63	
taacagaaa c	CCBABGCET TC 22	
210>	64	
211>	21	
212>	DNA	
213>	Artificial sequence	
220>		
223>	Porward primer for PCR amplification of AthURIQUE SSLP marker	
	in Arabidopsis thatiana subspecies	

e4005	64	
aggcaaatgt	coatttratt g	ı
<210>	65	
<211>	20	
<212>	DNA	
<213>	Artificial sequence	
<220>		
<2233	Reverse primer for PCR amplification of AthUBIQUE SSLP marks in Arabidopsis thaliana subspecies	ìΥ
<400>	6 5	
acgacatggc	agatttctcc 20	5
<210>	F.C.	
<310>	66 21	
<212>	DNA	
<213>	Artificial sequence	
(213)	As circles sageones	
<22D>	<i>;</i>	
<223>	Forward primer for PCR amplification of NGA172 SSLP marker i Arabidopsis thaliana subspecies	ת.
<400>	66	
agetgettee	ttatagegte c 2:	i
<210>	67	
<211>	19	
<212>	DNA	
<213>	Artificial sequence	
<220>		
<223>	Reverse primer for PCR amplification of NGA172 SSUP marker i	
	Arabidopsis thaliana subspecies	, 13
< 400>	67	
catoogaatg	ccattgtc)
<210>	68	
<211>	21	
<212>	DNA	
<213>	Artificial sequence	
<230>		

```
< 223 %
                 Forward primer for PCR amplification of NGA136 SSLP marker in
                 Arabidopsis chaliana subapecies
   <400×
                 58
   gasasascgc tactttcgtg g
                                                                              21
   <210>
                59
   <211>
                22
   <212>
                DNA
   <213>
                Artificial sequence
   <220%
   <223>
                Reverse primer for PCR amplification of NGA126 SSLP marker in
                Anabidopsis thalisma subspecies
   <400>
                б9
  caagagcaat atcaagagca gc
                                                                             32
  <210×
                70
  <211>
                23
  <212>
                DNA
  <213>
                Arcificial sequence
  <220×
               Forward primer for PCR amplification of NGA162 SSLP marker in
  <223>
               Arabidopsis thaliana subspecies
  <400>
               70
  catgemattt geatergagg
                                                                            20
  <210>
               71
  <211>
               22
  <212>
               DNA
  <213>
               Artificial sequence
 <220>
 <223>
               Reverse primer for PCR amplification of NGA162 SSLP marker in
               Arabidopsis thaliana subspecies
 <400>
               71
 ctctgtcact cttttcctct gg
                                                                            22
 <210>
               72
 <211>
               21
 <212>
              ANG
<213>
              Artificial sequence
```

WO 99/19492 PCT/EP98/06977

<220>		
<323>	Forward primer for PCR amplification of NGA6 SSLP marker Arabidopsis thaliana subspecies	in
<400>	72	
tggatttctt	cctetettea e	21
<210>	73	
<311>	21	
<212>	DNA	
<213>	Artificial sequence	
<220>		
<223>	Reverse primer for PCR amplification of NGA6 SSLP marker Arabidopsis thaliana subspecies	in
<400>	73	
atggagaagc	ttacactgat c	31

<210>	74 26	
<2125	DNA	
<213>	Artificial sequence	
<2285	The second secon	
<2235	Forward primer for PCR amplification of NGA12 SSLP marker Arabidopsis thaliana subspecies	in
<400>	74	
aatgttgtcc	teceteete	20
	•	
<310>	75	
<211>	22	
<312>	DNA	
<213>	Artificial sequence	
<220>		
<223>	Reverse primer for PCR amplification of NGA12 SSLP marker Arabidopsis thalians subspecies	in
<400>	75	
tgatgctctc		22

<210>	
<210>	76
<211><212>	21
	DNA
<\$13>	Artificial sequence
<220>	
<223>	Forward primer for PCR amplification of NGA8 SSLP marker in Arabidopsis thaliana subspecies
<400>	76
gagggcasa	it ctttattteg g
<210>	77
<211>	22
<212>	DNA
<213>	Artificial sequence
<220>	
<223>	Paulane mui - a d
	Reverse primer for PCR amplification of NGA8 SSLP marker in Arabidopsis chalicum subspecies
<400>	77
tggetttegi	t ttatasacat cc 22
<210>	78
<211>	21
<375>	DNA
<223>	Artificial sequence
2205	•
(223>	Forward primer for PCR amplification of NGA1107 SSLP marker in Arabidopsis thaliana subspecies
400>	78
cgaaaaaac	aaaaaatcc a
310>	
310> 211>	79
211> 212>	21
	DNA
213>	Artificial sequence
220>	
223>	Reverse primer for PCR amplification of NGA1107 SSLP marker in Arabidopsis thaliana subspecies
100>	79

capesacea	acagaartag g		21
<210>	8 C		
<211>	21		
<212>	DNA		
<213>	Artificial sequence		
	wrethierar aedreuce		
<220>	•		
<223>	Forward primer for PCR amplification of NGA225 SSL Arabidopsis thaliana subspecies	P marker	in
<400>	30		
gaeatccaaa	teccagagag g		21
<210>	B1		
<211>	22		
<212>	DNA		
<213>	Artificial sequence	•	
<220>			
<223>	REVETER Origon for Dou		
	Reverse primer for PCR amplification of NGA225 SSLA Arabidopsis thaliana subspecies	P marker	in
<400>	81		
totocccact	agttttgtgt ce	.;	22
<210>	82		
<311>	19		
<212>	DNA		
<213>	Artificial sequence		
<220>			
<223>	Paramed mades of a same at the		
~ ~ ~ ~	Forward primer for PCR amplification of NGA249 SSLE Arabidopsia thaliana subspecies	? marker	in
<400>	82		
taccgtcaat	ttcatcgcc	נ	L 9
<210>	B3		
<211>	22		
<212s	DNA		
<213>	Artificial sequence		
<220>			
<223>	Reverse primer for PCR amplification of NGA249 SSLP Arabidopsis thaliana subspecies	marker	in
	Commence of the state of the st		

```
<400>
                  83
    ggatcoctaa otgtamaato oo
                                                                               22
   <210>
                 84
   <211>
                 22
   <212≥
                 DNA
   <213>
                 Artificial sequence
   <220>
   <223>
                Porward primer for PCR amplification of CA72 SSLP marker in
                Arabidopsis thaliana subspecies
   <4005
                84
  eatcocagte accasacacs ca
                                                                              22
  <2105
                85
  <3115
                20
  <212>
                DNA
  د215ء
               Armificial Requosce
  <220>
  <223>
               Reverse primer for PCR amplification of CA72 SSEP marker in
               Arabidopsis challana subspecies
 <4005
               85
 cccagtctaa ccacgaccac
                                                                            20
 <210>
               86
 <2115
              20
 <2125
              DNA
 <213>
              Artificial sequence
 <220>
 <2235
              Forward primer for PCR amplification of NGA151 SSLP marker in
              Arabidopsis thaliana subspecies
<400>
              86
gtttcgggaa gttttgctgg
                                                                           20
<210>
              87
<2115
              24
<212>
             DNA
<213>
             Artificial sequence
<220s
```

<2239	Reverse primer for PCR amplification Arabidopsis thaliana subspectes	of	NGA151	sslp	marker	in
<400>	87		•			
cagtotadaa	gogagagtat gatg	•				24
<310>	B3					
<211≥	22					
<212>	DNA .					
<213≥	Artificial sequence					
<220>	•					
<223>	Parward primar for han sentification				•	
~~~	Porward primer for PCR amplification Arabidopsis thaliana subspecies	OI	NGA106	SSLP	marker	in
<400>	88					
gttatggagt	ttctagggea eg					53
<210×	89					
<211>	20					
<212>	AND					
<213>	Artificial sequence					
<230>						
<223 ×	Reverse primer for PCR amplification Arabidopsis thelians subspecies	of	NGA106	SSLP	marker	in
<400>	B9					
tgccccattu	tgttcttctc					20
<210>	90					
<211>	20	•				
<212>	AKC					
<313>	Artificial sequence					
<220>						
<223>	Forward primer for PCR amplification	_4	MORIDA	COTA		4 m
	Arabidopais thalians subspecies	OL	NGALIS	23116	MRIXEL	111
<400>	90					
agagotacca	gatergatgg					20
-330:			400		٧	
<210>	91					
<211>	21					
<212><213>	DNA					
~~	Artificial sequence					

<220>	
<223>	Reverse primer for PCR amplification of NGA139 SSLP marker in Arabidopsis thaliana subspecies
<400>	91
ggtttcgtt	t cactatocag g
<210>	92
<211>	22
<212>	DNA
<213>	Artificial sequence
<220>	
<223>	Forward primer for PCR amplification of NGA76 SSLP marker in Arabidopsis thaliana subspecies
<400>	92
ggagaaat	Ecactotoca co
<210>	93
<211>	20
<212>	DNA
<213>	Artificial sequence
<220>	
<223>	Reverse primer for PCR amplification of NGA76 SSLP marker in Arabidopais thaliana subspecies
<400>	93
aggcatggga	gacatttacg 20
<210>	94
<211>	20
<212>	DNA
<213>	Artifícial sequence
<220×	
<223>	Porward primer for PCR amplification of ATHSO191 SSLP marker in Arabidopsis thalisms subspecies
<400>	94
ctccaccaat	Catgossatg *

	A.P.
<210>	95
<211>	<b>21</b>
c212>	DNA
<213>	Artificial sequence
c220>	
<223>	Reverse primer for PCR amplification of ATHSO191 SSLP marker
,,	in Arabidopsis theliana subspecies
	in wrapinoberg cuminand ambabacies
<400>	05
4400>	95
•	
cdardredse	ggagatggtc a 21
<210>	96
<211>	22
<212>	DNA
<213>	Artificial sequence
<220>	•
<223>	Forward primer for PCR amplification of NGA129 SSLP marker in
	Arabidopsis thaliana subspecies
	wide independent and adoption
<400>	96
44005	30
+ magaaaaa	ANA
rcaggaggaa	ctaaagtgag gg
<210>	on
	97
<211>	22
<212>	DNA
<213>	Artificial sequence
<230>	
<223>	Reverse primer for PCR amplification of NGA129 SSLP marker in
y.	Arabidopsis thaliana subspecies
<400>	97
	•
CACACTGRAG	atggtettga gg
<210>	98
<211>	8062
<212>	DNA
<213>	
~6437	Arabidopsis thaliana ecotype Columbia
. 200	
<220>	
<223>	Ganomic DNA sequence of AtMSR6
. = -	
<400>	97
ztttttggct	get acaata asggtatacg grittatgte atea tatan etatatata 60

aagaaasgaa ayatatatat tgttttttca tttatcaac aaaacaacaa gactttttt	120
ttacttttta cattggtcaa casaatacaa gatasacgac atcgeetaat catttcccaa	180
ttttaccott aagtttaada totagaadot totcoatott egoaageaca gootgattag	240
gaacagettt accattetea tatteetgaa etacetgagt ceteteattg atetgttteg	300
ccasateege tigigacate ticticieca alctegetet englateate aaceteacet	360
elgettteae acgatecate geogragget etgittette tiecagette teegigtiaa	420
traceggaar egeogtagat treceettt tgtregaace ggearegaat ttettaaceg	480
tttgaaccgc gacaccgttt ctcagagctg cgttaaccgc ttccggatcg cgtaggtett	540
ggctettttg ttttgatetg tggagaacta ctggtteeca gtettgtgtt actgcteetg	500
ggtatetget eggeategte gatgaattga gagaaaggaa caacgegaaa attttattaa	660
tetgageett gaaattgaga aacgatgaag atgaagaatg tegttgagag gattgtgata	720
tttatacata egaagategg tttetggaga attegateat exected attitegtet	780
ctggaacqtt cttagagatg attgacgacq tgtcattatc tgatttgcag ttaaccaatg	840
ctttttgggt tggattegtg gtacaccata ttatecgatt tggctcaatg gttttatata	900
aattoggoto toggttoggt tatgagtoat cattamaatt magetaacem magattotog	960
taaaatttat tteggtttea atteggater ettaetteea gaategaatt attegaaacr	1020
ggggtragee gaacegaata ceaatgeeeg attgaetege tggotagaaa gateeaaegg	1080
tatecsatas tagascatas atoggacogo, catossagoo toaaagagto aacagtoasc	1140
aaaaaagtt gagocotgag gagtatogtt toogotattt otacgaogea aggogaaaat	1200
ttttggeger aatotttooc coctttogaa tteteteage teaaaacate gittetetet	1260
esetetetet cacaattees aaasatgeag egeesgagat egsetttgte tetetteesa	1320
anaccenegg eggegactae gaagggtteg gttteeggeg atgetgetag eggeggggge	1380
ggcagcggag accacgattt aatgtgaagg aaggggatgc caaaggcgac gcttctgtac	1440
sttttgctgt ttcgaaatct gtcgatgagg ttagaggaac ggacactcca ccggagaagg	1500
teegegteg tgteetgeeg tetggattta ageeggetga ateegeeggt gatgettegt	1560
cotqttctc caatattaty cataagtttg taaaagtoga tgatogagat tgttctggag	1620
gaggtacta atettogatt ctcttaattt tcttatettt agetggaaga agaagatteg	7680

rgraactigt	racarregee	ggagagacte	tgattattge	scragatogt	tgtttacana	1740
titttcaggag	cogagaagat	gtigticege	tgaatgatec	atototatgt	atgaaggota	1800
atgatgttat	tectraattt	cgttccaaca	atggtaaaac	tcaagaaaga	aaccatgott	1860
ttagrttcag	tgggagaget	gaacttagat	cagtagaaga	tataggagta	gatggcgatg	1920
cteetggtee	agaaacacca	gggatgegte	cacgtoctte	tegettgaag	cgagttctgg	1980
aggatgaaat	gacttttaag	gaggataagg	tteetgtatt	ggactctaac	aaaaggctga	2040
aaatgctcca	ggatccggtt	tgtggagaga	agaaagaagt	ssacgasgga	accasatttg	2100
aatggettga	gtettetega	atcagggatg	ccsatagaag	Acgtoctgat	gatcecettt	2160
acgatagaaa	gaccttacac	ataccacctg	atgttttcaa	gamaatgtot	gcatcacasa	2220
agcaatatcg	gagtgttaag	agtgaatata	tggacattgt	gettttettt	aaagtggtta	2280
gtaactatta	atctagtgtt	caatroatct	cctcaatgcg	atttgttcac	ttacatotgt	2340
tracgitate	CtCttCtCag	gggaaatttt	atgagctgta	tgagctagat	gcqgaattag	2400
gtcacaagga	gcttgactgg	aagacgaeca	tgagtggtgt	gggaaaatgc	agacaggtaa	2460
attagttgaa	acaactggcc	tgottgaatt	attgtgtcta	taaattttga	caccaccttt	2520
tgtttcaggt.	tggtatetet	gaaagtggga	tagatgaggc	agtgcaaaag	ctattagete	2580
grąggtaagg	gaaccatcat	actttatgga	attegtttae	tgctacttcg	gctaggattt	2640
aagaaatgga	aatcacttca	agcatcutta	gtraggatec	tgagaactca	ggatgttttc	3700
ttättogtta	tataataagt	cttttcatca	aggagtaaca	<b>AACAAAAC</b> tt	gcacaatatt	2760
tgtgtgctca	ctggcaaggr	atatataccc	agctaacctt	tgctagttca	ctgtagtaac	2820
agttacggat	aacatatgtt	tacttgtatg	tggtaccetc	attittgtctc	tcatggaggc	2880
tttcaagcct	tgtgttgasa	ctggacagtt	acatațgott	ccaacagaaa	ctagcatgca	2940
gattcatatg	ctttcctatt	ctactsatta	tgtattgaca	cactcgttgt	ttcttttgaa	3000
agatataaag	ttggacgaat	cgagcagcta	gaaacatctg	accaagcaaa	agceagaggt	3060
gctaatactg	taagttttct	tggataggtc	aaggagagtg	ttgcagactg	tttttgatca	3120
tttcttttc	tgtacattac	tttcatgctg	taattaacte	aatggctatt	ctggtctgat	3180
tatcagataa	Ctccaaggaa	gctagttcag	gtattaacte	catcaacage	aagegaggga	3240
aacat.coooc	ctgatgecgt	COSTOTECTT	great <i>ass</i> ag	ACOSETECTES	tttackvatt	3300

tatettatea tetteaette atecameta-	
tatettatea tetteagtte atecatette tegaalastta enetettett taccastett	3360
ceatcanget gegthaagga titggaatta gaaaatcatt attegatget tegtethata	3420
igeaagaggt tecettgaaa agatetgitt aagattetti geaccigaaa aatteaatet	3480
ttttaagtga atcrectact ttcttacaat gatcatagtc tgcsattgca tgtcsagtsa	
tateatteet tgttactgca tecepetett tettaatgae cattgtetat gttgtgtttg	3540
totogtgtgd tggagaaaat gatagotgat ocaagotgta cattatoatg attaagtago	3600
tgercaggaa ttgeetttgg ttacattgee taatggtttg atgreaattt ttettetgaa	3660
totttattt agateaaat gaareess.	3720
tetttaettt agateaaaat ggagetadaa aagtgtteaa etgtgtatgg atttgetitt	3780
gttgaetgtg etgeettgag gttttgggtt gggtctatea gegatgatge atcatgtget	3840
getettiggag egitattgat geaggiaage aagtgtatte tgtatettat gigtaceatg	3900
tgacttcctg tgcstatatt tgggttgcag gaactaattc tgaatcacca tttggtatgt	3960
tttttreagg tttrteraaa ggaagtgtta tatgacagra aaggtaaact gettgtateg	4020
ccagtegett tgttaaacag aatttaaggt aaatgacact ggttaattta aagtgcatac	4080
atgitgaeat attgoagggo tatcaagaga agcacaaaag go <b>tctaagga</b> aatatacgot	4140
gacaggtace attteagtag gcaagetaae tgacaattta acegeteace gaatgatagg	4200
totottaaac attgotaatg tagatgatgt ttatgtttca atctaatagg gtotacggcg	
gtacagttgg ctccagtacc acaagtaatg ggggatacag atgctgctgg agttagaaat	4250
ataatagaat ctaacggata ctttaaaggt tcttctgaat catggaactg tgctgttgat	4320
ggtctatatg aatgtgatgt tgccttagt ggtgtaactg tgctgttgat	4380
ggtctaaatg aatgtgatgt tgreettagt getettggag agetaattaa teatetgtet	4440
Aggeratage tegteteget tegtetagett tegetetea catattage anaggaactt	4500
ttestasett acagtereta tetaetryca yeragasgat gtaettaage atggggatat	4560
ttttccatac caagtttaca ggggttgtct cagaattgat ggccagacga tggtasatct	4620
tgagatattt aacaataget gtgatggtgg teetteagge aagtgeatat ttetttttig	4680
araacticaa ctagagggca gacatagaag gaaaaattet.aataettegt acggatetec	4740
agtaagtaat agcogatitt tgtttacota tgtagggaco ttgtacaaat atottgataa	4800
ctgtyttagt ccaactgyta agogactett aaggaattgg atctgocate cactcamaga	
tgtagaaago atcaataaad ggcetgatgt agttgaagaa etcacggcaa actcagaaag	4860
	4920

tatgcaaatc	actggccagt	atctccacaa	acttecagae	ttagaaagac	tgctdggacg	4950
catcaagtet	agcgttcgat	catcagecte	tgtgttgcct	getettetgg	ggaaaaaagt	5040
gotgaaacaa	cgagtaagta	tcaatcacaa	gttttctgag	taatgrette	catgagtagt	5100
acaggactaa	aacattacgg	gtotagetaa	agactgttct	cettetttg	caatgtctgg	5160
ttattcatta	catttetett	aacttattgc	attgcagget	aaagcatttg	ggcaaattgt	5220
gaaagggttc	agzagtggaa	ttgatctgtt	gttggctcta	cagaaggaat	Cadatatgat	5280
gagtttgett	tatasactct	gtsaacttcc	tatattagta	ggaaaaagcg	ggctagagtt	5340
attotttot	caattcgaag	cagccataga	tagegaettt	ccaaattatc	aggtgccat	5400
ctatotttca	tacettacaa	caaaatgtct	gtcactactc	asagesatge	atatggctta	546D
gateteaaet	cacaccccga	ggatcctaaa	gggatttgct	tittatteet	aatgtttttg	5520
gatggtttga	ttatttcta	acttgaactt	attaatettg	taccagaacc	aagatgtgac	5580
a <b>ga</b> tgaaaac	gccgaaactc	tcacaatact	tatcgaactc	tttatogaaa	gagcaactes 🔑	5640
aEggtctgag	gtcattcaca	rcataagctg	cctagatgtc	ctgagatett	ttgcaatcgc 🖟	5700
agcsagtctc	tctgctggaa	gcatggccag	gcctgttatt	tttcccgaat	cagaagetae	5760
agatcagaat	cagaasacas	aagggccaat	acttaaaatc	caaggactat	ggcatccatt	5820
tgcagttgca	gccgatggtc	aattgeetgt	toogaatgat	atactecttg	gcgåggctag	SBBO
aagaagtägt	ggcagcattc	atecteggte	attgttactg	асдддассаа	acatdādcāā	5940
aaaatcaact	cttcttcgtg	caacatgtct	ggeegttate	tttgcccaag	tttgtatact	6000
cgttagetaa	ttactctatt	ctttgcaatc	agttcttcaa	catgaataat	aaattotgtt	6060
trergtetge	agcttggctg	ctacgtgccg	tgtgagtett	gegaaatete	cctcgtggat	6120
actatettea	caaggcttgg	cgcatctgat	agaatcatga	caggagagag	taagttttgt	6180
totdaaaata	ccaattcctc	gaactattta	ctcagatttt	gtctgattgg	acaaggtggt	6240
tttgctttt	tttaggtacc	tttttggtag	aatgcactga	gacagcgtca	gttcttcaga	6300
atgcaactca	ggattcacta	gtaatccttg	Acgaactggg	cagaggaact	agtactttcg	6360
atggatacge	cattgcatac	toggtaacet	getettetee	ttcaacttat	acttgttgat	6420
Caacraaaac	atgcaattca	ttttgctgaa	acttattgat	ttatatcagg	tttttcgtca	6490
cctggtagag	aaagttcaat	greggatget	ctttgcaaca	cattaccacc	ctctcaccaa	6540

ggaattogog totoaccoac gigicaccio gaaacacatg gotigogeat teaaatcaag	SEAA
atotgattat caaccacgig grigigatca agacctagig therigiaco giblaaccga	6600
gggagettgt eetgagaget aeggaettea agtggeaete aeggetggaa taccaaacea	
agragetigaa acageateag gracigetea agecatgaag agateaatta gagaaaactt	6720
caagtcaagt gagctaagat ctgagttoto aagtotgcat gaagactggc tcaagtcatt	6780
ggrgggtatt totogagtog occasages traces in	€840
ggrgggtatt totogagtog orcacaacaa tgodoocart ggrggaagatg actacgacac	6900
ttlgttttgc ttatggcatg agatcaaatc ctcttactgt gttcccaaat aaatggctat	6960
gacaraacac tatergaage tegitaagie tittgettet etgatgitta ticetettaa	7020
aaaatgotta tatatoaaaa aattgtttoo togattataa caagattata tatgtatotg	7080
toggtttage tatggtatat aatatatgta tgttcatgag attggtcaag agaaatacto	7140
acaaacagta tattaagaag gaaatatgtt tatgcattaa tttaagtttc aagataaact	7200
gcazataacc tegactadag tigcaaagae caaacacaaa tiacaaaact tataagaett	7260
aagitetgaa tteectaaaa eeaaaaaaa aaacagaaca tattttgttg catetacaaa	7320
caacacaaac ctacatagtt tataacttac tcatcactga gattaacatc agaatcattc	7380
tecatotett catottoact otratoatoa teaceaceae catgatgatt etectectet	7440
tracgtaace tageastete actetgaget etatesacas tetgettett etgesactee	7500
adatitioned gadaatrage totoatotte tecaseteet teattegete tetettacte	7560
ttetecatet teteataaa etteccaaac eteteaacag aateegecaa catettatae	
gaageagegt cattaacett ettecteteg tactesacet catcatecte atectectee	7620
tobteagaat caccaggact atecatcate teatcadace cattagactt atetadataa	7680
accitagigt tratasacar saartraret gaatcaacar cacaagetaa acciasatet	7740
gačitogogo amacacana cancatata a colonia de calcado calcado calcada a colonia de calcada	7800
gaditgggcg aaacacaaag caacatatec aacttattga aaaacgacca titacitgaa	7860
dotabacetg atticicade citaatetic tottitedat Acticoteti caagicatea	7920
abdattotec tacattgogt ctragattto todatcetta gotoccace cactttotra	7980
getachteat becasseete gsteereaaa eteettetae eesattgesk sascetater	8040
receasactt caageaacae aa	9062

### **PCT**

#### WORLD INTELLECTUAL PROPERTY ORGANIZATION International Bureau



### INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6: C12N 15/82, 15/29, C07K 14/415, C12N 15/10, 5/04, 5/14, A01H 5/00

 $\mathbf{A3}$ 

(11) International Publication Number:

WO 99/19492

(43) International Publicati n Date:

22 April 1999 (22.04.99)

(21) International Application Number:

PCT/EP98/06977

(22) International Filing Date:

9 October 1998 (09.10.98)

(30) Priority Data:

PO 9745

10 October 1997 (10.10.97)

ΑU

(71) Applicant (for all designated States except US): RHONE-POULENC AGRO [FR/FR]; 14/20, rue Pierre Baizet, F-69009 Lyon (FR).

(72) Inventors; and

(75) Inventors/Applicants (for US only): DOUTRIAUX. Marie-Pascale [FR/FR]; 64, route de Villebon, F-91160 Saulx les Chartreux (FR). BETZNER, Andreas, Stefan [AU/AU]; 40 Dallachy Place, Page, ACT 2614 (AU). FREYSSINET, Georges [FR/FR]; 21, rue de Nervieux, F-69450 Saint Cyr au Mont d'Or (FR). PEREZ, Pascal [FR/FR]; 17, chemin de la Pradelle, Varennes, F-63450 Chanonat (FR).

(74) Agent: GENIN, Patrick; Rhône-Poulenc Agro, DPI, 14/20, rue Pierre Baizet, F-69009 Lyon (FR).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

#### **Published**

With international search report.

Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

4

(88) Date of publication of the international search report. 24 June 1999 (24.06.99)

### (54) Title: METHODS FOR OBTAINING PLANT VARIETIES

#### (57) Abstract

An isolated and purified DNA molecule comprising a polynucleotide sequence encoding a polypeptide functionally involved in the DNA mismatch repair system of a plant.

## FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL AM AZ BA BB BE BF BG BJ BR	Albania Armenia Austria Australia Azerbaijan Bosnia and Herzegovina Barbados Belgium Burkina Faso Bulgaria Benin	ES FI FR GA GB GE GH GN GR HU IE	Spain Finland France Gabon United Kingdom Georgia Ghana Guinea Greece Hungary Ireland	LS LT LU LV MC MD MG MK	Lesotho Lithuania Luxembourg Latvia Monaco Republic of Moldova Madagascar The former Yugoslav Republic of Macedonia Mali	SI SK SN SZ TD TG TJ TM TM	Slovenia Slovakia Senegal Swaziland Chad Togo Tajjikistan Turkmenistan Turkey Trinidad and Tobago
CG CH CI CM CN CU CZ DE DK EE	Congo Switzerland Côte d'Ivoire Cameroon China Cuba Czech Republic Germany Denmark Estonia	KE KG KP KR KZ LC LI LK LR	Kenya Kyrgyzstan Democratic People's Republic of Korea Republic of Korea Republic of Korea Kazakstan Saint Lucia Liechtenstein Sri Lanka Liberia	NL NO NZ PL PT RO RU SD SE SG	Niger Netherlands Norway New Zealand Poland Portugal Romania Russian Federation Sudan Sweden Singapore	YN YU ZW	Viet Nam Yugoslavia Zimbabwe

Inte. onal Application No PCT/FP 98/06977

			/EP 98/06977
IPC 6	HFICATION OF SUBJECT MATTER C12N15/82 C12N15/29 C07K14, C12N5/14 A01H5/00	/415 C12N15/10	C12N5/04
According	to International Patent Classification (IPC) or to both national classification	ication and IPC	
B. FIELDS	SEARCHED		
Minimum d IPC 6	ocumentation searched (classification system followed by classification sy	ation symbols)	
Documenta	tion searched other than minimum documentation to the extent that	such documents are included in	the fields searched
Electronic o	lata base consulted during the international search (name of data t	ase and. where practical, search	terms used)
C. DOCUM	ENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the re	elevant passages	Relevant to claim No.
A	WO 95 15381 A (CHAPELLE ALBERT D JOHNS HOPKINS (US)) 8 June 1995 see the whole document	E ;UNIV	1-5, 33-35
A	WO 97 37011 A (SETRATECH S A R L RHONA HARRIET (GB); LOUIS EDWARD 9 October 1997 see abstract see the whole document	;BORTS JOHN (GB)	1-35
A	WO 90 07576 A (SETRATECH) 12 Jul see the whole document and speci 5, line 16-31, examples 3-4	_	1,7-27, 31,32
	· ·	-/ <del>-</del> -	
X Furth	er documents are listed in the continuation of box C.	X Patent family members	are listed in annex.
"A" documer consider of filing da "L" documer which is citation "O" documer other m"P" documer	nt which may throw doubts on priority claim(s) or s cited to establish the publication date of another or other special reason (as specified) nt referring to an oral disclosure, use, exhibition or	cited to understand the prin invention  "X" document of particular relevicannot be considered nove involve an inventive step will document of particular relevicannot be considered to involve an inventive step will document is combined with	ordict with the application but ciple or theory underlying the same; the claimed invention for cannot be considered to hen the document is taken alone ance; the claimed invention volve an inventive step when the one or more other such docularing obvious to a person skilled
	ctual completion of the international search	Date of mailing of the intern	
	6 April 1999	03/05/1999	
Name and m	ailing address of the ISA European Patent Office, P.B. 5818 Patentiaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo rd,	Authorized officer	
	Fax: (+31-70) 340-3016	Mateo Rosell	, A.M.

onal Application No C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT PCT/EP 98/06977 Citation of document, with indication where appropriate, of the relevant passages Relevant to claim No. WO 97 01634 A (ANGELETTI P IST RICHERCHE Α BIO ; JIRICNY JOSEF (IT); PALOMBO FABIO () 1.2 16 January 1997 see page 1, line 18 - page 2, line 28 see page 58 PROLLA T A ET AL: "MLH1, PMS1, AND MSH2 Α INTERACTIONS DURING THE INITATION OF DNA 1,2,4,5 MISMATCH REPAIR IN YEAST" SCIENCE. vol. 265, 19 August 1994, pages 1091-1093, cited in the application Α A. WATANABE ET AL., : "Genomic organization and expresssion of the human 1-6 MSH3 gene" GENOMICS. vol. 31, 1996, pages 311-318, XP002099967 see the whole document and specially Fig. 3 Α ACHARYA S ET AL: "hMSH2 forms specific mispair-binding complexes with hMSH3 and 1-6 PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA, no. 93, October 1996, pages 13629- 13629, XP002080159 cited in the application sequences from this paper deposited in GenBank under AC= U61981 (hMSH3), and U54777 and U73732-7 (hMSH6). see the whole document Α K. LIU ET AL., : "characterization of the mouse Rep-3 gene: sequence similarities to 1-6 bacterial and yeast mismatch-repair proteins" GENE, vol. 147, 1994, pages 169-177, XP002099968 see the whole document and specially Figure 2. I. IACCARINO ET AL., : "MSH6, a Α Saccharomyces cerevisiae protein that 1-6 binds to mismatches as a heterodimer with MSH2" CURRENT BIOLOGY. vol. 6, no. 4, April 1996, pages 484-486, XP002099969 sequence is deposited at GenBank under AC= Z47746. see the whole document -/--

1

Form PCT/ISA/210 (continuation of second sheet) (July 1992)

Inte one Application No
PCT/FP 98/06977

STIAN) DOCUMENTS OF STREET	PCT/EP 98/06977
passages	Relevant to claim No.
CORRADI A ET AL: "CDNA SEQUENCE, MAP, AND EXPRESSION OF THE MURINE HOMOLOG OF GTBP, ADNA MISMATCH REPAIR GENE" GENOMICS, vol. 36, no. 2, 1 September 1996, pages 288-295, XP000613761 see the whole document and specially Fig.1	1-6
DATABASE EMBL NUCLEOTIDE AND PROTEIN SEQUENCES,13 July 1998, XP002099368 HINXTON, GB AC= AJ007791. Arabidopsis thaliana mRNA for mismatch repair protein (MSH3) see abstract	1-6
DATABASE EMBL NUCLEOTIDE AND PROTEIN SEQUENCES,1 August 1998, XP002099369 HINXTON, GB AC= 065607. Arabidopsis thaliana. Putative mismacht DNA repair protein see abstract	1-5
DATABASE EMBL NUCLEOTIDE AND PROTEIN SEQUENCES, 16 October 1997, XP002099371 HINXTON, GB  AC= AF009657. Arabidopsis thalian MutS homolog 6-like protein mRNA. see abstract  -& K.M. CULLIGAN AND J.B. HAYS: "DNA mismatch repair in plants" PLANT PHYSIOLOGY, vol. 115, 1997, pages 833-839, XP002099372 see the whole document	1-6
DATABASE EMBL NUCLEOTIDE AND PROTEIN SEQUENCES,17 December 1998, XP002099373 HINXTON, GB AC= AJ131669. Triticum aestivum mRNA for MSH3 protein, MSH3 gene. see abstract	1-6
DATABASE EMBL NUCLEOTIDE AND PROTEIN SEQUENCES,12 October 1998, XP002099370 HINXTON, GB AC= AJ007792. Arabidopsis thaliana DNA mismatch repair protein, MSH6 gene. see abstract	1-6
	EXPRESSION OF THE MURINE HOMOLOG OF GTBP, ADNA MISMATCH REPAIR GENE" GENOMICS, vol. 36, no. 2, 1 September 1996, pages 288-295, XP000613761 see the whole document and specially Fig.1 DATABASE EMBL NUCLEOTIDE AND PROTEIN SEQUENCES, 13 July 1998, XP002099368 HINXTON, GB AC= AJ007791. Arabidopsis thaliana mRNA for mismatch repair protein (MSH3) see abstract DATABASE EMBL NUCLEOTIDE AND PROTEIN SEQUENCES, 1 August 1998, XP002099369 HINXTON, GB AC= 065607. Arabidopsis thaliana. Putative mismacht DNA repair protein see abstract DATABASE EMBL NUCLEOTIDE AND PROTEIN SEQUENCES, 16 October 1997, XP002099371 HINXTON, GB AC= AF009657. Arabidopsis thalian MutS homolog 6-like protein mRNA. see abstract & K.M. CULLIGAN AND J.B. HAYS: "DNA mismatch repair in plants" PLANT PHYSIOLOGY, vol. 115, 1997, pages 833-839, XP002099372 see the whole document DATABASE EMBL NUCLEOTIDE AND PROTEIN SEQUENCES, 17 December 1998, XP002099373 HINXTON, GB AC= AJ131669. Triticum aestivum mRNA for MSH3 protein, MSH3 gene. see abstract DATABASE EMBL NUCLEOTIDE AND PROTEIN SEQUENCES, 12 October 1998, XP002099370 HINXTON, GB AC= AJ007792. Arabidopsis thaliana DNA mismatch repair protein, MSH6 gene.

 $\mathbf{G}$ 

1

Form PCT/ISA/210 (continuation of second sheet) (July 1992)

Information on patent family members

PCT/EP 98/06977

Patent do	Cument		Publication	T			98/06977
Cited in sea			date	1.	Patent family member(s)		Publication date
WO 951		A	08÷06-1995 	EP JP US US		09 T 70 A	11-09-1996 30-06-1997 02-12-1997 17-11-1998
WO 9737	011	Α	09-10-1997	NON	 E		
WO 9007	·	A	12-07-1990	FR AU CA DE DE EP ES IE JP	264179 12751 483439 200654 6892417 6892417 044992 2077058 72469 4503601	9 T 0 A 9 A 4 D 4 T 3 A B B	20-07-1990 15-09-1995 01-08-1990 26-06-1990 12-10-1995 14-03-1996 09-10-1991 16-11-1995 09-04-1997 02-07-1992
WO 97016	34	A 	16-01-1997	IT AU	RM950434 6241296	A A	27-12-1996 30-01-1997